

orbit

ISE/UT's Magazine for Schools

VOL. 36, NO. 2, 2006

36:2

MAKING CLASSROOM ASSESSMENT WORK FOR STUDENT LEARNING

GUEST EDITORS:

Clay Lafleur

Lorna Earl

Steven Katz

OISE UT

www.orbitmagazine.ca



OISE/UT's Magazine for Schools

VOLUME 36, NUMBER 2, 2006

MAKING CLASSROOM ASSESSMENT WORK
FOR STUDENT LEARNING

Guest Editors:

Clay Laflleur, Lorna Earl, and Steven Katz

Orbit publishes theme issues on topics of concern and relevance for teachers and administrators working in schools and school systems, from junior kindergarten to the end of high school.

EDITOR

Heather Berkeley

ACTING ASSISTANT EDITOR

Dawn Martin

EDITORIAL BOARD

Barbara Bodkin, Continuing Education, OISE/UT
David Booth, Scholar in Residence, OISE/UT
Karl Dehl, Sociology and Equity Studies in
Education, OISE/UT

Patricia Goldblatt, Ontario College of Teachers
Kim Gordon, The Bishop Strachan School
Linda Grant, University of Ontario Institute
of Technology

Clay Laflleur, The Learning Partnership
Peter Lipman

Tait Luste, Peel District School Board
Kirk Mark, Toronto Catholic District School Board
Pat McAdie, Elementary Teachers' Federation of
Ontario

John Myers, Curriculum, Teaching and Learning,
OISE/UT

Patti Pappas, McGraw Hill Ryerson

David Roker, Academic Technology, OISE/UT
Carol Rolheiser, Associate Dean, OISE/UT
Paul Vanderhelm, Halton District School Board

ORBIT EDITORIAL OFFICE

OISE/UT
252 Bloor Street West
Toronto, Ontario M5S 1V6
Phone: 416-923-6641, ext. 2681
Fax: 416-975-1925

E-mail: orbitmagazine@oise.utoronto.ca
Web site: www.orbitmagazine.ca

LAYOUT: Tracy Choy, BTT Communications

PRINTED AND BOUND IN CANADA

BTT Communications

ORBIT SALES AND CIRCULATION

OISE/UT
252 Bloor St. West
Toronto, Ontario M5S 1V6
Phone: 416-923-6641, ext. 2077
Fax: 416-975-1925
E-mail: orbitmagazine@oise.utoronto.ca

Second Class Mail
Registration Number 4511
ISSN 0030-4433

Canadian Publications Mail Product Agreement
No. 638528
GST # R108162

TABLE OF CONTENTS

- 1 **EDITORIAL**
Classroom Assessment for Learning
Clay Laflleur, Lorna Earl, and
Steven Katz
- 2 **SECTION 1**
ASSESSMENT AND LEARNING
Introduction
Steven Katz
- 2 **Assessment for Learning –**
Why, What, and How?
Dylan Wiliam
- 7 **Putting the Right Accent on Learning:**
The Role of Regulation in Formative
Assessment
Dany Laveault
- 10 **Critical Learning Instructional Path**
(CLIP): Assessment for Learning in
Action
Carmel Crévola, Peter Hill, and
Michael Fullan
- 16 **SECTION 2**
ASSESSMENT FOR LEARNING IN ACTION
Introduction
Lorna Earl
- 16 **Feedback for Learning**
Ruth Sutton and Pauline Clarke
- 20 **Collaborating with Students in the**
Assessment Process
Damian Cooper
- 24 **Grading Group Work: Dilemmas of**
Policy, Research, and Practice
John Myers
- 27 **Assessment Tracking Boards:**
A Focus for Professional Dialogue
Mary Lou McKinley, Kathy Witherow,
and Angela Puma
- 30 **SECTION 3**
RETHINKING CLASSROOM ASSESSMENT
Introduction
Clay Laflleur
- 30 **Standardized Testing,**
Classroom Assessment, Teachers,
and Teacher Unions
Pat McAdie and Ruth Dawson
- 34 **Reducing Bias in Classroom Assessment**
and Evaluation
Louis Volante
- 37 **All Aboard! The Land of Formative**
Assessment Awaits
Carolynne Muncer
- 39 **What Do We Want Rubrics to Do for Us**
and Our Students?
Judy Arter
- 45 **Creating Coherent Formative and**
Summative Assessment Practices
Lorrie Sheppard

PULL OUT

Orbit Professional Development:
Learning Activities

Carol Rolheiser and Barb Bower

SPECIAL THANKS

To board member, Carol Rolheiser, for playing the role of
Critical Friend in the production of this issue.

ILLUSTRATIONS

The images on the cover and throughout this issue are those
of Kathy Di Santo, a recent graduate of OISE/UT's teacher
education program. Kathy donated her work "Rubrics
Cube" to OISE/UT in 2005 following the annual teacher
education art display Kidspeak. The images on the cube are
photographs of student self-portraits. Thanks to the artist,
her students, the organizers of Kidspeak, especially Terry
Reeves, and Larry Swartz who brought this artwork to our
attention.

ARTIST'S STATEMENT

"Games and youth culture are intimately related. This, in
addition to my understanding of the adolescent learner as a
multi-dimensional individual who struggles with issues of
identity on a daily basis, is the imaginative force behind my
work. I wanted to create a piece that communicated these
messages through visual interpretation alone and through
the process of engaging with the cubes. This is an interactive
piece. The inspiration comes from one of my students who
was obsessed with his Rubik's cube. This three-dimensional
puzzle is both fascinating and frustrating. I wanted to make
tangible the feelings that teenagers experience on a normal
basis and felt that I could achieve this through an interactive
approach, without compromising the integrity of their lives.
The images on the cube are photographs of my Grade 11
collage self-portraits that were part of a series of portraiture
work."

OUR MISTAKE

It was Drew Langston (a teacher with the Peel District School
Board and an OISE/UT grad) who drew the images for "Drawn
to the Page," the graphic story published in the previous *Orbit*
issue, *Literacy Revisited*, Vol. 36, No. 1, 2006.

Classroom Assessment *for* Learning

Assessment matters. If a commitment to the success of every child is the heart of our moral contract, then assessment that is integrally connected to learning is the life-blood that nourishes it.

We have moved beyond a mere acceptance of assessment of what students know and are able to do as the only way to assess students' learning. Our thinking, planning, and action are now guided by a belief that assessment must enrich and be an integral part of the total learning process. Assessment *for* learning and assessment *as* learning are more encompassing and complementary ways of thinking about assessment to improve instruction and student achievement.

Without question, quality assessment is hard work. Over the last several years we have learned a great deal about how to improve our assessment practices. Assessment supports learning when it is directly linked to the goals of instruction. And, as the authors emphasize, daily and specific assessment practices which respect the diversity and background of each student make a difference.

Improving our assessment practices and existing structures is still very much a political, emotional, and technical undertaking. In the past, assessment was often a taken-for-granted and unscrutinized activity. Pausing to critically reflect on our current practices can move us forward. By developing the ability to explore and be curious about our own experiences and actions, we open up the possibilities of purposeful learning for ourselves. In this sense, reflection becomes the foundation for new learning and becomes a door opener to deeper understanding and ultimately targeted action that benefits our students.

Assessment is complex. It directs and adapts instruction. It guides capacity building. It moves students forward. It empowers students to take charge of their own learning. It informs students, teachers, and parents about the impact of their efforts. The use of assessment for and as learning has a bigger impact on student achievement than any other intervention. When the goals of assessment are clear and precise, when assessment practices are fair, and when assessment is a daily habit, then learning becomes a purposeful and meaningful activity. While assessment is about closing the gap between a student's actual performance and the targeted level of achievement, it is also about many other issues such as control, choice, engagement, and meaning.

Placing the learner at the centre of learning at any given moment is crucial. As the authors in this issue articulate, assessment is about helping students grow. For this to occur, we must continually interrogate and reflect on our practice and the assumptions that guide our actions. We must take an active role in ensuring that assessment and learning are intimately connected. Further, instruction must be precise and focused; it must be routinely available on a regular basis for all students. Assessment is about where students are, where they need to go, and how best to get there.

Focused teaching requires teachers to have precise and continuously updated information—from the beginning as well as during the learning process. Assessment is about making learning

visible and authentic. Each student's individual needs must be respected as they move along established patterns of development. Assessment is about collaborative communication between teachers, students, and parents. It is an optimal flow experience that is nurtured daily.

As the authors demonstrate, our professional knowledge about assessment is continually improving. We know a lot more than we did. Our task, should we choose to accept it, is to use this knowledge to improve student learning—now.

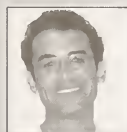
— GUEST EDITORS



CLAY LAFLEUR, Director of Policy and Research at The Learning Partnership, is also an instructor in the Department of Theory and Policy Studies in Education at OISE/UT. Clay has taught in elementary and secondary schools as well as colleges and universities, both in Canada and Australia. His teaching and research interests include policy development, knowledge transfer, innovations, leadership, educational change, assessment, program evaluation, and inquiry.



LORNA EARL, Director, Aporia Consulting Ltd., is a recently retired Associate Professor in the Theory and Policy Studies Department and Head of the International Centre for Educational Change at OISE/UT. She is currently Researcher In Residence with the Ontario Ministry of Education. Lorna has worked for over 20 years in schools and school boards and, as a leader in the field of assessment and evaluation, has been involved in consultation, research, evaluation, and staff development with teachers' organizations, ministries of education, school boards, and charitable foundations in Canada, England, Australia, Europe, and the United States.



STEVEN KATZ, Senior Lecturer in Human Development and Applied Psychology at OISE/UT, is also the coordinator of the Psychology of Learning and Development initial teacher education component. He has received the Governor General's medal for excellence in his field and has recently served as co-principal investigator for the external evaluation of England's Networked Learning Communities program. Steven is a director of Aporia Consulting Ltd., and co-author, with Lorna Earl, of *Leading Schools in a Data-Rich World* (Corwin Press).

Assessment and Learning



The relationship between classroom assessment and learning is undoubtedly complicated. Fortunately, ongoing work has served to disentangle some of the complexities. We know, for example, that classroom assessment is imbued with a multi-purpose character. At times, it must certify proficiency of students; at other times, it must direct teachers in the next steps of their instructional sequences; and, at still other times, it must cultivate the capacity within students to take charge of their own learning.

In *Rethinking Classroom Assessment with Purpose in Mind* (2006, www.wncp.ca), Lorna Earl and I outline these three distinct purposes of classroom assessment as:

1. assessment of learning,
2. assessment for learning, and
3. assessment as learning.

ASSESSMENT FOR LEARNING:

Why, What, and How?

■ Dylan Wiliam

DYLAN WILIAM sees that teacher quality is the single most important variable in student progress and he highlights assessment for learning as the practice whereby teachers use evidence about student achievement to adapt instruction to better meet learning needs.

Education matters. It matters for individuals and it matters for societies. For the individual, greater levels of education are associated with greater control over one's life, greater economic

wealth, and a longer and healthier life (OECD, 2003, p. 33). For societies, increasing levels of educational achievement result in lower criminal justice costs, lower costs of state-provided health care (Schweinhardt et al., 2005), and increased economic growth. Quantifying the amount of extra economic growth inevitably involves some pretty heroic assumptions. Nevertheless, Hanushek (2004) has calculated that if we could increase student achievement by one standard deviation—equivalent to about 15 percentage points on a typical test where the scores range from around 20 per cent to around 80 per cent—then, in 30 years, the increased growth in the economy would more than pay for the entire cost of K–12 education.

We also note the importance and validity of each in the education enterprise and argue for the need to be able to differentiate among the purposes of classroom assessment in the service of high-leverage, intentional practice. Indeed, even the Assessment Reform Group in the U.K., the collaborative synonymous with promoting the assessment for learning agenda, has recently worked on the differentiation challenge by explicitly considering the role of teachers in assessment of learning.

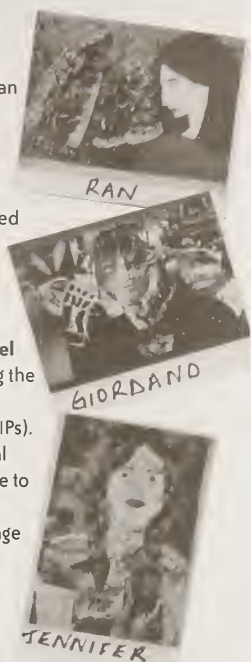
Each of the articles in Section 1 considers the relationship between assessment and learning:

- **Dylan Wiliam** sees teacher quality as the single most important variable in student progress and he highlights assessment for learning as the practice whereby teachers use evidence about student achievement to adapt instruction to better meet learning needs.
- **Dany Laveault** connects the psychological notion of “regulation” to formative assessment.

He explains the role that assessment can play in both the learning of self regulation as well as in the self regulation of learning. Laveault argues that a better understanding of the role of regulation can contribute to improved contextualization and differentiation in formative assessment practices by, for example, taking students’ developmental levels into account.

- **Carmel Crévola, Peter Hill, and Michael Fullan** extend this latter point by noting the place of assessment for learning within Critical Learning Instructional Paths (CLIPs). An essential part of developing a critical learning path, they explain, is being able to make a direct link between a student’s results on a set of assessments, their stage of development, and specific teaching strategies relevant to that stage of development.

—Steven Katz



Where’s the Solution?

Around the world, many approaches to school improvement have been proposed, but most have been relatively ineffective because they have failed to embrace a lasting and inevitable truth: the single most important variable in the amount of progress that a student makes at school is the quality of the teacher.

This has not been obvious because for many years we lacked the right kinds of data-sets. There is a huge variability in the levels of achievement of students in different schools, and simplistic analyses suggested that these differences must be attributable to the schools. However, it rapidly became clear that one of the major reasons for these differences in the “outputs” of schools was that the schools served very different communities. Analyses that took into account the socio-economic status (SES) of students in different schools painted a very different picture—it was SES, rather than school quality that was the major determinant of educational achievement. Still others attributed the differences to the students themselves; their prior achievement for example.

However, as more and more data-sets began to consider the “value added” by the school—the difference between what the students knew when they started at that school and what they knew when they left—it became clear that the quality of the

teacher was the most significant variable. If you get one of the best teachers, you will learn in six months what an average teacher will teach you in a year, and if you get one of the worst teachers, the same amount of progress will take you two years—there is a fourfold difference between the productivity of the best and worst teachers (Hanushek, 2004; Wiliam & Thompson, in press).

Improving education is therefore, in effect, a labour-force issue with two solutions. One is to replace existing teachers with better ones. The problem with this is that there is no evidence that there are better teachers queuing up to get in, but who can’t get jobs (Darling-Hammond et al., 2005). Nor is there that much evidence that better pay would have a substantial impact on teacher quality (there are many good reasons for paying teachers more, but the supply-and-demand argument is not one of them). The other solution is to invest in the teachers that we have—what my colleague Marnie Thompson calls the “love the one you’re with” strategy.

Twenty years ago, this would have resulted in a very gloomy prognosis, since there was little evidence that it was possible to improve the quality of a teacher’s practice, except very slowly, and at great cost (Fullan, 1991). Many people therefore concluded that the situation was hopeless, and that policy makers should, instead, focus on solutions that did not require quality teaching

(such as producing "teacher-proof" textbooks, so that it didn't matter how good the teacher was).

However, the problem with this argument is that we can conclude very little from the failure of years of teacher professional development to improve teaching because, for the vast majority of teachers, the kind of PD they received flew in the face of what the research says about what makes for effective PD. That research says that PD needs to be continuing, rather than in the form of "one-shot" inputs (Cohen & Hill, 1998); it needs to be tailored to the local circumstances in which teachers work (Cobb et al., 2003); it needs to be directly related to the content of the subjects they teach (Supovitz, 2001); and it needs to involve teachers in active and collective participation (Garet, Birman, Porter, Desimone, & Herman, 1999).

So much for the "process" of teacher PD. What about the "content" of teacher PD? Does it matter what teachers focus on, as long as they are engaged in the processes listed above? It turns out that it does. It turns out that teachers can spend a lot of time developing their thinking, without changing their practice very much, and from the research it appears that one particular focus—the use of assessment for learning—has a bigger impact on student achievement than any other.

Assessment for Learning

For over 20 years, Paul Black and I have been researching the role that assessment can play in raising levels of student achievement. In particular, we wanted to find out if using assessment to support learning, rather than just to measure its results, could improve students' achievement, even when such achievement is measured in the form of state-mandated tests. In reviewing 250 studies from around the world, published between 1987 and 1998, we found that a focus by teachers on assessment for learning, as opposed to assessment of learning, produced substantial increases in students' achievement—typically doubling the rate of learning (Black & Wiliam, 1998). Since the studies also revealed that day-to-day classroom assessment was relatively rare, we felt that considerable improvements would result from supporting teachers in developing this aspect of their practice.

One of the things that we have struggled with is what, exactly, is assessment for learning? Many people have proposed definitions, all of which differ slightly, but it seems that the "big idea" is that evidence about student achievement is used to adapt instruction to better meet learning needs (see Wiliam & Thompson, in press). The key strategies involved in this are:

- clarifying and sharing learning intentions and success criteria with learners;
- engineering effective classroom discussions, activities, and tasks that elicit evidence of student achievement;
- providing feedback that moves learners on;
- activating students as owners of their own learning; and
- activating students as instructional resources for one another.

Building Capacity for Assessment for Learning

KEY PARAMETERS

1. Accountability

Write out an action plan; give monthly reports on what was tried, and the results.

2. Support

Agree on an agenda for peer observation; what should count as evidence with reference to the action plan.

3. Choice

Use techniques within a framework of accountability, always being able to answer, "What's formative about that?"

4. Flexibility

Modify techniques as needed, keeping the research evidence in mind.

5. Gradualism

Proceed slowly to effect faster change; implement no more than three changes at one time into your practice.

There is ample evidence that focusing on any one of these five strategies is highly effective in raising student achievement (Wiliam, in press). Together they provide the highest leverage focus that we know for improving schools, classrooms, and the achievement of students. The problem is that knowing what needs to be done is one thing. Doing it is quite another.

Putting It into Practice

We know that "one size fits all" does not work in teacher professional development. What might be exactly the right thing to do in one situation might be exactly the wrong thing to do in a similar, but different, situation. Expert teachers do not use general, all-purpose approaches to solve problems, but rather generate solutions that take advantage of specific details in the challenges they face (Berliner, 1994). That is why, "what works" is not the right question in education; everything works somewhere, and nothing works everywhere. The right question is "under what conditions does this work?" That is why if we are serious about developing assessment for learning, we must help each teacher find her or his own way of doing this.

The opposite of one size fits all is to allow each teacher to choose what they want to do. This may be attractive, but it is not

rigorous. Teachers may choose to change things that have no impact on student achievement. The trick is, therefore, to generate teacher professional development that is open enough to allow each teacher to adapt the new ideas into their own practice, but structured enough so that the adaptation does not produce a “lethal mutation” (Pellegrino, 2006) that renders the innovation impotent.

Over the last three years, I and my colleagues at the Learning and Teaching Research Center at the Educational Testing Service in Princeton, NJ, have been working with many groups of teachers on the best ways to implement assessment for learning. Some approaches have been extremely successful, and others less so, but even from our failures, we have usually been able to learn something (and often more from our own failures than from our successes). As a result of reflection on both our successes and failures, we have been able to identify five elements that increase the successful implementation of assessment for learning: accountability, support, choice, flexibility, and gradualism.

1. ACCOUNTABILITY

Most professionals involved in teacher development will have had the experience of generating considerable enthusiasm for, and commitment to, change during a summer workshop, only to find that all their good intentions seem to be erased by the demands of the new school year. That is why we suggest that teachers should make a commitment in writing about what they are going to change about their practice (the action plan) and then be held accountable by colleagues at monthly meetings for making those changes in their practice. Each month, every teacher describes what he or she tried and how it went. Teachers have told us that having to face their colleagues and “deliver” on the promises they made the previous month helps them move their “change” task to the top of their in-tray.

2. SUPPORT

The other side of the coin of accountability is support. We have found that at the monthly meetings teachers are able to offer each other advice when the planned changes are not going well, and the fact that these meetings are with groups of peers, rather than an expert teaching a novice, appears to be particularly beneficial. It must be borne in mind that, ultimately, implementing assessment for learning involves changes in day-to-day and even minute-to-minute classroom practice. That is why we also recommend that teachers engage in peer observation. To clearly distinguish these observations from those routinely carried out to manage performance, these observations should be done by genuine peers rather than those in a hierarchical relationship. Another important requirement is that the teacher being observed must set the agenda for the observation and spell out for the observer what should count as evidence, by reference to her or his action plan. By defining the observer’s role, both in terms of what is to be looked for and

what counts as evidence, the observer’s own prejudices are minimized, and the difference between this and supervisory observation is emphasized.

3. CHOICE

Teachers often describe the process of changing their practice as “scary.” Implementing assessment for learning in particular makes many teachers feel as if they are being asked to “give up control” of their classrooms. However, teachers responsible for choosing what they will change about their practice feel empowered, especially when they can choose among techniques those that appeal to them. This choice lies, however, within a framework of accountability. While teachers are free to choose what they change, they are accountable for changing something. They are also accountable for showing their peers how their innovations are consistent with the principles of assessment for learning. One especially powerful way to support this is for peers to ask each teacher, “What’s formative about that?” in order to emphasize the need for the changes in practice to be focused on using assessment evidence to adapt instruction to meet student needs.

4. FLEXIBILITY

A technique that works for one teacher may not work for another, but may do so after some modification. Teachers need to be encouraged to modify techniques to make them work in their classrooms, but they also need to understand enough of the research evidence so that the changes they make do not render the innovation ineffective.

5. GRADUALISM

Asking teachers to make wholesale changes in their practice is a little like asking a golfer to change her swing during a tournament. Teachers have to maintain the fluency of their classroom routines, while at the same time disrupting them. The action plans that teachers develop should specify a small number of changes—ideally two or three—that they will make in their teaching. In our experience, teachers who try to change more than three things in their practice at the same time are never successful. They try to do too much, and their classroom descends into chaos, and as a result they revert to what they know how to do. In other words, going slower produces faster real change.

Conclusion

Our experience is that, equipped with some basic ideas about assessment for learning, teachers can support each other in making radical improvements in their students’ learning, provided each teacher is responsible for her or his own development targets. It is perfectly OK to hold teachers accountable for making changes in their practice, provided that:

- the teachers have a choice in how they put the basic principles of assessment for learning into effect;

- they are not pushed to make changes faster than they can incorporate them into their normal practice; and
- they have the support and accountability that a like-minded group of peers can provide.

References

- Berliner, D.C. (1994). Expertise: The wonder of exemplary performances. In J.N. Mangieri & C.C. Block (Eds.), *Creating powerful thinking in teachers and students: Diverse perspectives* (pp. 161–186). Fort Worth, TX: Harcourt Brace College.
- Black, P. J., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles Policy and Practice*, 5(1), 7–73.
- Cobb, P., McClain, K., Lamberg, T.d.S., & Dean, C. (2003). Situating teachers' instructional practices in the institutional setting of the school and district. *Educational Researcher*, 32(6), 13–24.
- Cohen, D.K., & Hill, H.C. (1998). *State policy and classroom performance: Mathematics reform in California*. Philadelphia, PA: University of Pennsylvania Consortium for Policy Research in Education.
- Darling-Hammond, L., Holtzman, D.J., Gatlin, S.J., & Vasquez Heilig, J. (2005). Does teacher preparation matter? Evidence about teacher certification, teach for America, and teacher effectiveness. *Education Policy Analysis Archives*, 13(42).
- Fullan, M. (1991). *The new meaning of educational change*. London: Cassell.
- Garet, M.S., Birman, B.F., Porter, A.C., Desimone, L., & Herman, R. (1999). *Designing effective professional development: Lessons from the Eisenhower Program*. Washington, DC: US Department of Education.
- Hanushek, E.A. (2004). *Some simple analytics of school quality* (NBER working paper no. W10229). Washington, DC: National Bureau of Economic Research.
- Organisation for Economic Cooperation and Development. (2003). *Education at a glance*. Paris, France: Organisation for Economic Cooperation and Development.
- Pellegrino, J.W. (2006, Mar 17). Personal communication.
- Schweinhardt, L.J., Montie, J., Xiang, Z., Barnett, W.S., Belfield, C.R., & Nores, M. (2005). *Lifetime effects: The High/Scope Perry preschool study through age 40*. Ypsilanti, MI: High/Scope Educational Research Foundation.
- Supovitz, J.A. (2001). Translating teaching practice into improved student achievement. In S.H. Fuhrman (Ed.), *From the capitol to the classroom: Standards-based reform in the States* (Vol. Part 2, pp. 81–98). Chicago, IL: University of Chicago Press.
- Wiliam, D. (in press). Keeping learning on track: formative assessment and the regulation of learning. In F.K. Lester, Jr. (Ed.), *Second handbook of mathematics teaching and learning*. Greenwich, CT: Information Age Publishing.
- Wiliam, D., & Thompson, M. (in press). Integrating assessment with instruction: What will it take to make it work? In C.A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning*. Mahwah, NJ: Lawrence Erlbaum Associates.



DYLAN WILIAM is Deputy Director of the Institute of Education in London. After teaching in urban schools in London for seven years, he spent 19 years at King's College, London, where he was Dean of the School of Education, and Assistant Principal of the College. From 2003–06, he directed the Learning and Teaching Research Center at the Educational Testing Service in Princeton, NJ.



Common Threads:
From Canada to South Africa
Combating HIV/AIDS Together

Classroom resource Teaching Materials Integrated DVD and CD-ROM Lesson Plans Evaluation Rubrics

The Ontario Secondary School Teachers' Federation has created high quality, classroom-ready resources on combatting HIV/AIDS in South Africa and Canada. It is multimedia in nature, engages students and instructors alike and is designed for a variety of subjects, including: civics, science, health and social sciences.

CONTENTS:

program DVD, 200+ pages of hands-on lesson modules, supporting print materials and resources, CD-ROM and interactive website access.

PRICE:

\$50 — Complete package.
\$30 — Program DVD and CD-ROM.

ORDER:

Online at www.osstf.on.ca and select Common Threads or e-mail bellisd@osstf.on.ca.
Telephone: 416.751.8300 or 1.800.267.7867



Canadian
International
Development
Agency

cope
1998-01

PUTTING THE RIGHT ACCENT ON LÉARNING:

The Role of Regulation in Formative Assessment

■ Dany Laveault

DANY LAVEAULT connects the psychological notion of *régulation* to formative assessment. He explains the role that assessment can play in both the learning of self regulation and in the self regulation of learning.

There is no simple answer to the question of what aspect of formative assessment best supports student learning. Is it more efficient assessment practices or improved instructional methods that makes this such a powerful assessment methodology? Part of the difficulty originates from the variety of formative assessment designs and the different conceptual frameworks that are used to develop them. Some of these frameworks result from different scientific traditions. Among French-speaking countries (France, Belgium, Switzerland, and Canada), formative assessment has evolved quite differently since its initial conception by Bloom (1968). Though the purpose remains the same in both the English and French traditions, some important distinctions have started to appear, especially around the notion of *régulation* in French.

The word "regulation" is present in the English literature but it is rarely encountered in relation to formative assessment. In the French literature, however, it is present to a much larger extent. The concept itself owes as much to developments in the French scientific tradition (such as Piaget's genetic epistemology [1975/1985]), as it does to developments in the English tradition (such as Bandura's socio-cognitive theory [1997] and Zimmerman's model of self regulation [2000; 2001]). The concept also borrows extensively from the cognitive sciences and theories of systems.

Definition of Regulation

Whichever scientific tradition we refer to, the concept of regulation involves three components:

- a goal or target;
- feedback; and
- a form of control.

That is why it so readily applies to formative assessment. Stiggins, Arter, Chappuis, and Chappuis (2004) implicitly refer to

the notion of regulation when they identify assessment for learning as addressing three main questions:

1. Where am I going? (target or goal);
2. Where am I now? (feedback); and
3. How can I close the gap? (between where I am now and where I am going).

Black and Wiliam's (1998) feedback system also implies the notion of regulation. The target or goal is created by the "gap" between data on the student's actual level and data on the reference level. The other two components of Black and Wiliam's system involve a mechanism for comparing levels to monitor the gap reduction (feedback) and a mechanism to alter the gap (remedial action).

Despite such global similarities, the domain of *régulation* in the French research has been considerably widened and goes far beyond the basic definition (Allal et al., 2005). As *régulation* acquired more significance in the French literature, important conceptual distinctions started to appear. A distinction was initially made by Allal (1988) among three modalities of regulation:

1. *Interactive regulation* occurs when formative assessment is based on the interactions of the student with the other components of the instructional activity—that is, with the teacher, with other students, and/or with material allowing self regulated learning.
2. *Retroactive regulation* occurs when formative assessment is conducted after completion of a phase of teaching and allows identification of the instructional objectives attained or not attained by each student. It corresponds to the notion of remediation present in the initial conception defined by Bloom.
3. *Proactive regulation* occurs when different sources of information allow the preparation of new instructional activities designed to take into account differences among

The more control and choice allowed to students, the less likely they are to blame their failure on external factors (teachers, subject matter, or activity).

Regulation of learning thus involves much more than the regulation of an activity through cognitive and metacognitive processes. It involves regulation of social as well as motivational processes of learning.

students. It is concerned with instructional differentiation to insure enrichment and consolidation according to student needs, rather than focusing on remediation of learning difficulties.

While Allal (1988) makes a distinction among regulations occurring before, after, or throughout an activity of learning, Laveault (2004) identifies three imbedded "layers" of regulations occurring within the regulation of a learning activity:

1. *Regulation of the interaction between the student and the activity itself.* At this level, regulation will occur to monitor progress towards the learning target and take action to move the student closer to it.
2. *Within the learning activity regulation, regulation may occur at the level of the interaction among agents or "actors" (student, teacher, student's peers).* At this level, regulation will occur to develop a shared teacher and student understanding of their respective goals. It may involve "negotiation" of goals or standards, the selection of a method of feedback, or appropriate remedial actions. Regulation occurring among agents at the feedback level may lead to taking the necessary actions to determine whether the students are receiving appropriate support from the teacher or peers at different stages of learning.
3. *Within the regulation among agents, regulation may also bear on the interaction between each agent and a role (tutor, tutored, assessor, assessed).* For instance, a student may be called to self assess or assess a peer. He may also be asked to tutor a peer or be tutored by a peer. At this level, regulation may occur among students to determine who will be the tutor and who will be tutored, what the goals of each one will be, and how better tutoring strategies may be used.

Functional and Dysfunctional Aspects of Regulation

Regulation is also quite powerful to describe effective as well as ineffective practices of formative assessment. In the same manner as it allows teachers to act upon a child's problem through the careful observation of the child's systematic errors, the study of formative assessment's dysfunctional regulations may help diagnose inappropriate use of feedback, goal setting, or remediation. Inappropriate regulations lead to inappropriate or

insufficient control over action, and consequently, no learning occurs under such conditions. Baumeister, Heatherton, and Tice's (1994) book on regulation failures summarized these problems under the notions of overregulation, misregulation, and underregulation:

- *Overregulation* will occur when too much feedback occurs from too many sources.
- *Underregulation* will occur when students are not fully aware of an activity's goal or of the teacher's assessment criteria.
- *Misregulation* will occur whenever feedback misses the point such as when it bears on the student's general aptitude rather than on the way a task was appropriately or inappropriately done.

The study of regulation failures is just as relevant as the study of successful regulations to account for the functional aspects of formative assessment (Laveault, 2006). A better understanding of the role of regulation may contribute to improved contextualization and differentiation in formative assessment designs. For instance, one important way to tailor formative assessment to students' needs and characteristics is to take into account the students' developmental level of cognitive development. Students at different stages of development are not capable of the same kind of regulations (Piaget, 1975/1985), nor can they regulate the same "objects."

Regulation, Self Regulation, and Self Assessment

Assessment plays an important role in the *learning of self regulation* as well as in the *self regulation of learning*. There are two main reasons for this. First, when self assessing, students learn as much on the task as they learn about themselves. Second, the regulation of a learning activity cannot proceed without some form of feedback loop where learners continuously compare their progress towards an achievement goal, which is an assessment in itself.

The more control and choice allowed to students, the less likely they are to blame their failure on external factors (teachers, subject matter, or activity). On the contrary, such students will feel more responsible for their actions. The same rule applies to self assessment. Allowing more student control through self assessment may influence students' beliefs systems, such as causal attributions and self efficacy perceptions (Bandura, 1997). Providing students more effective control on self assessments through clearer definition of goals and standards, for instance, is likely to increase the student's self efficacy perception and resilience.

Increased control over assessment by students is given progressively through appropriate regulations at the activity, agent, or role levels. The model of self regulation may be used to establish different degrees of students', peers', or teachers' involvement in the goal, feedback, or action component of self assessment. A whole family of self assessment procedures may

be defined within such a model, varying from total student involvement in goal setting, feedback, and remediation to total involvement by teachers in all three components, in which case no self assessment is implied (Laveault, 2004).

For instance, at the beginning of a learning sequence, goals and standards may be set up mostly by the teacher. With increased mastery, students may be progressively required to define additional performance criteria by themselves and to set higher targets of achievement. Students may also be invited to participate in some form of peer assessment or elaborate self assessment. As learning progresses, they may also be given more autonomy in seeking feedback by themselves through the assistance of their peers or by referring to appropriate classroom resources.

Conclusion

Formative assessment and self assessment are not dichotomous functions. Assessment may be more or less formative. Self assessment may be more or less under students' control. The concept of regulation allows us to characterize these on a continuum where different regulations contribute to make appropriate adjustments, depending on the students' characteristics and the learning context. Such designs help to make them work better by incorporating relevant aspects of theories of goals, theories of feedback, and models of remediation. Regulation of learning thus involves much more than the regulation of an activity through cognitive and metacognitive processes. It involves regulation of social as well as motivational processes of learning.

The concept of regulation has helped redefine formative assessment as an object of study in the French literature (Allal & Mottier Lopez, 2005; Allal, 2006). It has been especially influenced by socio-constructivist views and has retained the following principles:

1. Formative assessment must be supported by a valid and coherent design. The notion of regulation enables researchers and teachers to diversify formative assessment practices and to link learning theories with relevant assessment models to make learning happen.
2. Formative assessment designs must be adapted to take into account individual differences as well as the learning context. Student development intervenes in two ways:
 - in the choice of regulations teachers may use to make learning happen and,
 - in the capacity of the student to effect the appropriate regulations to adapt and learn.
3. Self assessment, as an important component of formative assessment, must be introduced gradually with appropriate and timely phasing in of student control and the phasing out of teacher control on different aspects of the regulation of a learning activity, whether it is about setting goals, accessing feedback, or taking appropriate actions.

References

- Allal, L. (2006). Régulations des apprentissages: Orientations conceptuelles pour la recherche et la pratique en éducation. In L. Allal & L. Mottier Lopez (Eds.), *Régulation des apprentissages en situation scolaire et en formation*. Bruxelles: De Boeck.
- Allal, L. & Mottier Lopez, L. (2005). FA of learning: A review of publications in French. *Formative Assessment: Improving learning in secondary classrooms*, pp. 241–264. Paris: OEDC.
- Allal, L. (1988). Vers un élargissement de la pédagogie de maîtrise : Processus de régulation interactive, rétroactive et proactive. In M. Huberman (Ed.), *Assurer la réussite des apprentissages scolaires? Les propositions de la pédagogie de la maîtrise*. Delachaux et Niestlé, Neuchâtel, pp. 86–126.
- Bandura, A. (1997). *Self-efficacy. The exercise of control*. New York: W.H. Freeman & Co.
- Baumeister, R.F., Heatherton, T.F., & Tice, D.M. (1994). *Losing control: How and why people fail at self regulation*. San Diego: Academic Press.
- Black, P. & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7–74.
- Bloom, B.S. (1968). Learning for mastery. *Evaluation Comment*, 1(2), 1–12.
- Laveault, D. (2006). De la régulation au réglage : Etude des dispositifs d'évaluation favorisant l'autorégulation des apprentissages. In L. Allal & L. Mottier Lopez (Eds.), *Régulation des apprentissages en situation scolaire et en formation*. Bruxelles: De Boeck.
- Laveault, D. (2004). Interactions entre formation et évaluation: de la régulation entre acteurs et leurs rôles à la régulation de l'apprentissage. *Mesure et évaluation en éducation*, 27(1), 51–67.
- Piaget, J. (1985). *The equilibration of cognitive structures: The central problem of intellectual development*. Translated by T. Brown & K.J. Thampy. Chicago: University of Chicago Press.
- Piaget, J. (1975). *L'équilibration des structures cognitives: Problème central du développement*. Paris: Presses universitaires de France.
- Stiggins, R.J., Arter, J.A., Chappuis, J. & Chappuis, S. (2004). *Classroom assessment for student learning. Doing it right—using it well*. Portland, OR: Assessment Training Institute.
- Zimmerman, B.J. (2000). Attaining self regulation. A social cognitive perspective. In M. Boekaerts, P.R. Pintrich & M. Zeidner (Eds.), *Handbook of self regulation* (pp. 13–35). San Diego: Academic Press.
- Zimmerman, B.J. (2001). Theories of self regulated learning and academic achievement: An overview and analysis. In B.J. Zimmerman & D.H. Schunk (Eds.), *Self regulated learning and academic achievement, 2nd edition*. Mahwah, NJ: LEA.



DANY LAVEAULT specializes in measurement and evaluation and is a full professor at the Faculty of Education, University of Ottawa. From 1993–2001, he was the director of the international publication *Mesure et évaluation en éducation*. In 1995, he was awarded the Benoît-Poulin prize for outstanding contributions to the field of educational evaluation.

CRITICAL LEARNING INSTRUCTIONAL PATH:

Assessment *for* Learning in Action

■ *Carmel Crévola, Peter Hill, and Michael Fullan*

An essential part of developing a critical learning path, explain CARMEL CRÉVOLA, PETER HILL, and MICHAEL FULLAN, is being able to make a direct link between a student's results on a set of assessments, their stage of development, and specific teaching strategies relevant to that stage of development.

A decade of experience has led us to propose what we see as the *breakthrough* that will transform classroom instruction and lead to quantum improvements in literacy outcomes. Not 75 per cent success (Ontario's current target for strong student achievement in literacy and numeracy), but 90 per cent or more students being literate. Because we are talking about all classrooms, this means large-scale reform. We propose a framework for doing this, a framework based on concrete work we are engaged in at all three levels of what we call the tri-level reform solution—the school (including every classroom) and community; the district; and the state or government levels (see Fullan, Hill & Crévola, 2006).

Breakthrough Components

A breakthrough will be achieved when virtually all students are served well by the public education system. The model, which we propose for achieving breakthrough results, is shown in Figure 1.

We do not go into detail in this article—the important point is that if we aspire to reach and engage all students we will need a system that is powerful enough to do the task. Models of differentiated instruction by themselves will not mobilize the whole system.

Personalization or individualization is what advanced instructional systems aspire to. Personalization puts the learner at the centre and attempts to provide an education that is tailored to the students' learning and motivational needs at any given moment.

Individualized instruction must be effective for every student, which means it must be "precise" without falling into what we call the "prescription trap." To be precise is to get something right; to prescribe is to lay down rigid rules. Precision is about addressing the unique needs of individuals, in a dynamic fashion on a daily basis. It is also about harnessing the power of "assessment for learning" so that individually tailored data are available and used on a daily basis.

In this article we focus on the school and classroom levels of the tri-level model and on the individualization of instruction using the Critical Learning Instructional Path (CLIP) model.

Assessment *for* Learning

Assessment *for* learning is the process of seeking and interpreting evidence for use by teachers and their students to decide where the students are in their learning, where they need to go, and how best to get there (Black & William, 1998; Sadler, 1989; Stiggins, 2004). This requires that teachers are provided with the tools necessary for effectively and efficiently gathering, analyzing, and making sense of the data.

There are key principles of assessment *for* learning that have guided our work with districts and schools—assessment *for* learning should:

1. be recognized as central to classroom practice,
2. be part of effective planning of teaching and learning,
3. focus on how students learn,
4. be regarded as a key professional skill of all teachers,

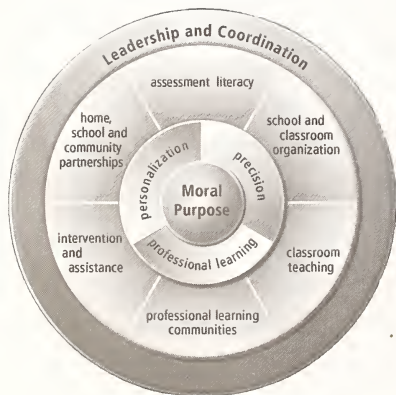


Figure 1

5. promote commitment to learning goals and shared understanding of the criteria by which students are being assessed,
6. provide constructive guidance about how to improve,
7. develop learners' capacity for self assessment so they can become reflective and self managing, and
8. recognize the full range of achievements of all learners.

Critical Learning Instructional Paths (CLIPs)

Given what we know about the problems facing systems trying to achieve 90 per cent literacy success, we need data-driven systems that focus on the learning needs of *all* students and teachers. The solution must be simultaneously efficient and effective. In particular, such a system must include four essential components:

1. A set of powerful and aligned assessment tools tied to the learning objectives of each lesson;
2. A method to capture assessment for learning data in a way that is not time-consuming, and impact instructional decisions immediately;
3. A means of using the assessment information on each student to design and implement personalized instruction; and
4. A built-in means of monitoring and managing learning, of testing what works, and of systematically improving the effectiveness of classroom instruction.

We have developed Critical Learning Instructional Paths (CLIPs) which are devices for bringing expert knowledge to bear on the detailed decisions that every teacher must make day-to-day and minute-to-minute in order to meet the individual needs of all students. We place teachers at the centre of this process as the experts. The core feature of this work requires *professional learning*, or learning in context on a daily basis, for all teachers.

A CLIP must be able to guide and monitor the learning and teaching on a day-to-day basis: it is therefore detailed in nature. Without detail, we cannot achieve precision. Whereas much traditional curricula focuses on end-points of instruction, a CLIP is about the actual route to be taken by the average learner, including the detours or loops followed by significant numbers of students who, for one reason or another, can be expected to take alternate pathways at certain points.

Mapping the Learning Journey

A helpful metaphor in thinking about a Critical Learning Instructional Path is the notion of a journey—but with different starting points for different individuals and with some individuals needing to take departures from the main pathway.

Start with a *description of the terrain* as a means of clarifying exactly what it is that students are doing when they embark on the journey of learning to read. These descriptors are “big picture” ideas that might form part of this initial step of clarifying the nature of the journeys taken by beginning readers.

Critical Learning Instructional Paths (CLIPs) 12-Step Process

- Step 1. Defining the terrain—big picture overview.
- Step 2. Key stages—mapping the journey.
- Step 3. Objectives—medium-term goals for instruction.
- Step 4. Indicators—specific and comprehensive short-term outcomes.
- Step 5. Pre-assessment—starting points for instruction.
- Step 6. Assessments—short and aligned.
- Step 7. Student learning profile—summaries of students' starting-, mid-, and end-points.
- Step 8. Focus sheet—planning, assessing, and evaluating “on the run.”
- Step 9. Instructional strategies matrix—aligning curriculum intent, assessment information, and instructional strategies.
- Step 10. Instructional strategies and grouping practices—a core of powerful strategies for use in whole- and small-group instruction.
- Step 11. Ongoing monitoring—focus sheets, student learning profiles, indicators.
- Step 12. Post-assessment—end-points of instruction.

Identify the key stages in the journey taken by learners. Returning to our early literacy example, in working with schools, we have highlighted six developmental stages of reading that typify the paths taken by young learners. Although we present these stages as a developmental continuum, we realize they do not necessarily form an invariant sequence. It is not uncommon to see students exhibiting a range of behaviours that span a number of stages, depending on the text that they are reading.

Outcomes	K	1	2	3
1. Oral Language	All	All	At Risk	At Risk
2. Benchmark Text Level/ Comprehension	All	All	All	All
3. Fluency	All	All	All	All
4. Concepts About Print	All	All	At Risk	At Risk
5. Phonemic Awareness	All	All	At Risk	At Risk
6. Letter Identification	All	All	At Risk	At Risk
7. Phonics	All	All	All	At Risk
8. Word Knowledge	All	All	At Risk	At Risk
9. Vocabulary	All	All	All	All

Figure 2. Example of a Literacy Assessment Regime: Grades K–3

The stages are:

1. pre-emergent
2. early emergent
3. emergent
4. beginning
5. transitional
6. established

Identify medium-term objectives or outcome statements and more detailed, short-term *indicators of progress*. The outcome statements provide medium-term goals for instruction—for example, what the typical student will have achieved by the end of the school year. The indicators of progress are short-term outcome statements that enable the teacher to trace the steps made by the learner as he or she moves from being a novice to having partial understanding or mastery to acquiring competence. They provide *feedback to teachers* on the effectiveness of their instruction and specific *instructional foci* for their daily lessons. They also form the basis for *feedback to students* to enable them to “self monitor” their learning, to evaluate their performance, and to know what constitutes an improved performance.

What should these indicators look like? Sometimes they will resemble and have the tight specificity of behavioural objectives, and will be amenable to temporal sequencing. This would apply, for example, to indicators of progress in acquiring letter-sound knowledge. Sometimes indicators will be of a kind that cannot be expressed as either present or not present, that require the use of metacriteria in arriving at overall judgments of quality, and that are not hierarchical in the sense that one logically precedes or follows another. This characterization would apply, for example, to indicators of progress in comprehending interpretive meanings of texts.

Mapping and Monitoring the Learning

Having mapped the CLIP followed by students in the early years as they learn to read, the next task is to design a system for measuring and monitoring their progress with reference to the key stages and the indicators. Focused teaching requires that teachers have precise and continuously updated information on students’ starting points and on their progress along the way.

The first step in measuring and monitoring learning is to specify the key outcomes that will be assessed throughout the CLIP. Next, a schedule of pre- and post-assessment must be established in order to measure the beginning and end points of students within each school year and with reference to the key stages of the journey as captured in the outcome statements for each stage (See Figure 2).

Few schools adopt an assessment regime that involves pre-assessment, and from experience we know how difficult it is to persuade schools to do so. An effective assessment regime needs to provide information on students’ instructional needs as well as acting as a “first alert” for those students who need direct intervention. Once these needs are established, teachers can more easily group students and provide focused

1	<i>That car in the yard used to be my Dad's.</i>	
2	<i>The boy on the horse is holding the reins.</i>	
3	<i>Over the holidays Mum bought me some sneakers.</i>	
4	<i>These are the pilots who fly the big planes.</i>	
5	<i>The girl made a snowman up on the mountain.</i>	
Subtotal		

Figure 3. Segment of oral language assessment

Stage Assessment	Pre Emergent	Early Emergent	Emergent	Beginning	Transitional	Established
Oral Language*	0-4	5-7	8-12	13-15		
Benchmark Text Level*	Not on text	A-B	B-D	D-G	G-K	K-S
Fluency*	N/A	20 WPM Level B text	30 WPM Level D text	40 WPM Level G text	40 WPM Level I text	90/110 WPM Level M/N text
Concepts About Print*	0-6	7-14	15-18			
Phonemic Awareness*	0-15	16-35	36-52			
Letter Identification	n-nn	nn-nn	nn-nn	nn-nn	nn-nn	nn-nn
Phonics	n-nn	nn-nn	nn-nn	nn-nn	nn-nn	nn-nn
Work Knowledge*	0-5	6-10	11-15	16-20		
Vocabulary	n-nn	nn-nn	nn-nn	nn-nn	nn-nn	nn-nn
Small-Group Instructional Strategies	- Read To - Oral Language - Language Experience - Shared Reading	- Read To - Oral Language - Language Experience - Shared Reading - Guided Reading	- Oral Language - Language Experience - Shared Reading - Guided Reading	- Oral Language - Language Experience - Guided Reading	Guided Reading	- Guided Reading - Reciprocal Teaching

* Keyed to instruments published in Crévola and Vineis (2004b)

Figure 4. Instructional Strategies Matrix

teaching that is targeted at the appropriate level. Focused teaching means ensuring that students are operating in their zone of proximal development, where competence and challenge are aligned. The hallmark of the expert teacher is that he/she operates in a focused instructional mode consistently, on a day-to-day basis, for all students. At the end of the year, information gained by using the same assessment tools allows measurement of the gains made during the year.

An essential requirement for making this do-able is a set of quick, easy-to-administer assessment tools. In recent years, thanks to the increased attention policy-makers have paid to early reading, a wide range of assessment tools have been developed for use in classrooms and schools. Figure 3 shows a segment of an oral language assessment for use with K-3 students (Crévola & Vineis, 2004a; 2004b). The instrument provides a means of assessing the structures of oral English that students understand and can control. As such, it is a

measure of receptive language. It consists of a series of sentences of increasing structural complexity.

A score of one point is awarded for each sentence repeated correctly in every detail. Altogether, a student is presented with 15 sentences and five sentence types. This particular assessment takes approximately five minutes to administer and yet it reveals a powerhouse of information to generate discussion and instructional exploration.

Using the Data to Drive Instruction

In addition to a method for capturing formative assessment data in a way that is not overwhelmingly time-consuming, teachers need a method of analyzing the data automatically and converting it into information that is powerful enough to drive instructional decisions. Finally, attention must be given to the design of procedures for monitoring on a daily basis student

progress with reference to the indicators. These all form part of the instructional regime of the classroom.

The first step is to generate summaries of students' starting points through the creation of Student Learning Profiles. These enable teachers to summarize all the assessment data generated through the pre-testing, to identify strengths and weaknesses, and to determine the stage of (reading) development of each student. Once this is done, teachers can effectively group their students to tailor the instruction in both whole class and small group learning settings (Crévola & Vineis, 2004b).

Being able to make a direct link between a student's results on a set of assessments, their stage of development, and specific teaching strategies relevant to that stage of development is an essential part of developing a CLIP. It makes the whole process of diagnosing students' strengths and weaknesses, and the planning of instruction explicit, and as such, amenable to ongoing improvement and refinement. Figure 4 gives an example of an Instructional Strategies Matrix.

It is clear that implementing a CLIP for early literacy depends very much on the effective use within classes of small instructional groups and a variety of powerful instructional strategies to provide focused teaching. This is where the Focus Sheet comes into play. Both the planning of the instruction and the monitoring of student progress are facilitated through the use of a Focus Sheet (Crévola & Vineis, 2004b; Fullan, Hill, & Crévola, 2006). The Focus Sheet ensures that the planning of instruction, the recording of student progress, and the evaluation of teaching are a single, seamless process.

Conclusion

CLIPs are devices for bringing expert knowledge to bear on the detailed decisions that every classroom teacher makes every day. The good news is educators know a great deal about early literacy: more than enough to construct a typical learning pathway and to specify appropriate instructional strategies. While all this expert knowledge is available, it is not currently in the hands of every teacher. CLIPs represent breakthrough thinking because they provide schools and teachers with powerful and easy-to-use tools to manage instruction in the most expert way possible.

References

- Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139–148.
- Crévola, C., & Vineis, M. (2004a). *Building essential literacy with Bookshop: A research-based reading program*. New York: Mondo Publishing.
- Crévola, C., & Vineis, M. (2004b). *Managing assessment and instruction: Grade 1*. New York: Mondo Publishing.
- Fullan, M., Hill, P., & Crévola, C. (2006). *Breakthrough*. Thousand Oaks, CA: Corwin Press; and Toronto: Ontario Principals' Council.
- Sadler, D.R. (1989). Formative assessment in the design of instructional systems. *Instructional Science*, 18, pp. 119–144.
- Stiggins, R. (2004). New assessment beliefs for a new school mission. *Phi Delta Kappan*, 86(1), 22–27.



Victoria, Australia.

CARMEL CRÉVOLA is an independent international literacy consultant, author, and researcher who works extensively in Australia, Canada, the U.K., and the U.S. She has 22 years of K–9 classroom teaching experience and was a school principal in



Education, University of Melbourne, and head of the state school system in Victoria, Australia.

PETER HILL is Secretary General of the Hong Kong Examinations and Assessment Authority. He was formerly Director of Research at the National Center on Education and the Economy, Washington D.C., Deputy Dean of the Faculty of



MICHAEL FULLAN is the former Dean of OISE/UT. Recognized as an international authority on educational reform, Michael is engaged in training, consulting, and evaluating change projects around the world. He led the evaluation team which conducted the assessment of the National Literacy and Numeracy Strategy in England from 1998–2003. In April 2004, Michael was appointed Special Advisor to the Premier and Minister of Education in Ontario.



OPEN the door to
research-based
teaching and learning.

ADVANCE your
career and enrich your
education experiences
through our
M.Ed. Degree program
in curriculum studies
and teacher
development.

Invest in your career with a Master of Education Degree
Course-only M.Ed., full or part-time studies

Possible THEMES to explore:

Literacy, Science, Technology in Education, Holistic
Education & Arts Education, Math, Social Studies & Citizenship,
Principles of Elementary Education, Urban Schools, Induction

Offered by:
Department of Curriculum, Teaching and Learning (CTL)
Ontario Institute for Studies in Education of the
University of Toronto

for information visit www1.oise.utoronto.ca/depts/ctl

*Celebrating 100 Years of Education
at the University of Toronto 1907-2007*



Assessment for Learning in Action



Assessment for learning occurs throughout the learning process. It is designed to make each student's understanding visible, so that teachers can decide what they can do to help students' progress. Students learn in individual and idiosyncratic ways, yet, at the same time, there are predictable patterns of connections and preconceptions that some students may experience as they move along the continuum from emergent to proficient.

In assessment for learning, teachers use assessment as an investigative tool to find out as much as they can about what their students know and can do, and what confusions, preconceptions, or gaps they might have. The wide variety of information that teachers collect about their students' learning processes provides the basis for determining what they need to do next to move student learning forward. It provides the basis for giving students descriptive feedback and deciding on groupings, instructional

Feedback for Learning

■ *Ruth Sutton and Pauline Clarke*

RUTH SUTTON and **PAULINE CLARKE** discuss the "feedback for learning" process used in Winnipeg inner city schools to help teachers think and work "smarter," structure learning experiences that fully engage the learner, and most of all, provide the steps for feedback to their students.

Feedback for Learning is an initiative to help teachers set targets for their students; to create independent student learners; and in the process, raise students' performance levels. It helps

students recognize their success and talk about it. The initiative has been developed in the Inner City District of Winnipeg School Division which serves a very diverse student demographic and comprises 21 schools—Nursery to Senior 4 and adults*.

Creating Effective and Meaningful Learning Settings

In the early 1990s, the Inner City District focused on the core business of school improvement—helping students learn through agreed principles of learning and principles of assessment. The main focus for all of our schools is to create

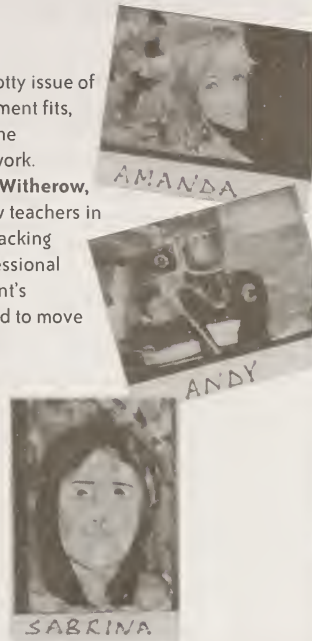
strategies, and resources (Earl & Katz, *Leading Schools in a Data Rich World: Harnessing Data for School Improvement*, 2006).

The articles in Section 2 give vivid images of different ways that teachers use assessment to unlock each student's learning:

- **Ruth Sutton** and **Pauline Clarke** discuss the "feedback for learning" process that they use in Winnipeg inner city schools to help teachers think and work "smarter," to structure learning experiences that fully engage the learner, and most of all, to provide the steps for feedback to their students.
- **Damian Cooper** puts students at the centre of the assessment process and describes how teachers play the role of coaches to help students become self assessors who become expert at understanding their own learning.

- **John Myers** explores the knotty issue of group work, and how assessment fits, with particular attention to the dilemmas of grading group work.
- **Mary Lou McKinley**, **Kathy Witherow**, and **Angela Puma** show how teachers in Ontario schools are using "tracking walls" as a stimulus for professional discussion about each student's learning, and what is required to move the learning forward.

—Lorna Earl



the most effective and meaningful learning settings for children. We challenge our staff to think about their practice and offer them the opportunity to talk about it in detail and in depth.

Our work on learning and assessment was informed by our successful experiences with the Inner City Arts Training Program. This showed us that in order to have sustainable change in teachers' practice we had to provide ongoing opportunities for learning, including trying new strategies, followed by reflection and discussion with peers.

Staff attend three professional development days during the year and then apply the newly-learned strategies to their work in classrooms. Teachers keep reflective journals and return to the next session ready to share their experiences with their colleagues. Throughout the year the program leaders visit the participants' classrooms to provide feedback in the "real" setting. Using our consultants as program leaders in this way had a positive impact on teachers changing their practice.

Pursuing Sustainable Change

Experience has taught us that we did not want to rely on any process that was not sustainable within existing resources. We couldn't rely on short-term external funding because such

funding usually disappears. The pursuit of sustainable change is time consuming, but the slow pace is a necessary investment in long-term growth.

After working for three years (1995–98) on our assessment principles and their implications, including trying to determine appropriate expectations from the provincial curricula, we came across research published by Paul Black and Dylan Wiliam from King's College, London, which summarized all the global findings from 1988 to 1998 on the link between assessment and improved learning and student achievement. Their findings confirmed our growing understanding that teacher feedback to students is the critical piece in children's learning in schools.

From then on, the Inner City District focused on incorporating Black and Wiliam's research findings into our classrooms. This research also crystallized the design of our Feedback for Learning initiative with its primary focus on setting goals for creating independent student learners.

The design of the Feedback for Learning project recognized the central role of school leaders as the main change agents. Consultants would be used to support the work, as in the Inner City Arts program, but ultimately changes in learning and teaching will have to be sustained by school teachers and leaders working together long after the consultants have moved on.

**Editors' note: Nursery is the year prior to Kindergarten, Senior 1=Grade 9, Senior 4=Grade 12.*

1. Continued Commitment of School Leaders

We have continuing conversations on many aspects of children's learning but use Feedback for Learning as an underpinning to all of them. The school leaders are themselves focusing on observing and gathering evidence of children's learning and talk together about the nature of "evidence." The "common language" of Feedback For Learning is used by our school leaders and staff; we speak the same language and share the same understandings.

2. Structure for Support

Clusters (seven clusters—three schools each) have been established whereby the administration and staff of these clusters meet, challenge each other's thinking, and share ideas. The learning support teachers appointed by the district are a key piece of this work and provide support to each cluster.

3. External Supports

We continue to have excellent support and input from inside the district in the person of Kathy Collis, and from outside through

Thompson Owens, Caren Cameron, and Ruth Sutton. All of these consultants are invited at key times to refresh the purpose and strategies of Feedback for Learning, and to help schools induct teachers new to the work.

4. Superintendent's Expectations and Support for the Climate of Learning

The superintendent has provided the framework, encouragement, and support to implement and continue this work. Her expectation has always been that children's learning is the priority for our schools even though many other issues can impede on this focus. She supports the risk-taking and creativity necessary to do this work, but also recognizes that perseverance is critical. The pace of effective change can be slow.

5. Joint Planning and Flexibility

We developed a plan which we review annually. The group of schools set the priorities for this work and made the choice to use existing funds. The flexible use of resources, such as staffing and budgets, help to make the work sustainable.

Parallel to the Inner City District focus on feedback to students, the Winnipeg School Division was developing a system-wide assessment process for all children Nursery—Grade 6—the Comprehensive Assessment Program (CAP). This process focused on English, mathematics, and French, and on children's hearing and social-emotional skills. The methodology was based on research indicating that teachers need to observe how children approach the assessment tasks, as well as assessing what the children achieve. Teachers were trained to gather information at the start of each school year to give them a view of the child's levels of achievement, the level of support they need, and their learning patterns.

This Comprehensive Assessment Program reinforced the critical importance of teachers working side-by-side with children to better understand how children learn so that teachers can provide relevant, appropriate programs. The first full year of the program (2000) coincided with the first year of Feedback for Learning, which required all of us to understand how summative and formative assessments can support each other although they may at first appear to pull in opposite directions.

By 1999, experience in the Inner City District had illuminated some key requirements needed to achieve the sustainability we wished for. We knew that teachers need opportunity to reflect deeply on their work, to be supported by skilled help, and that a slow pace of change in the early stages would help to foster the necessary roots. We also knew that working with volunteers would help the best classroom practice to emerge in the early stages. Before Feedback for Learning's official start in 2000, six of the 21 schools volunteered to be part of the first year. Within those schools, teachers were asked if they wished to be involved in the "first wave." In one small school, all the teachers joined in, but in others we had between four and eight teachers in each of the schools. To support them, one of the existing Inner City support teachers was asked to devote half his time to this work, and the district sought another consultant to bring both expertise and long-term involvement to the enterprise.

The choice of this person was influenced by prior experience: we knew that we needed someone who would respect our starting point and work with us. Technical expertise would be helpful, but the prime characteristic would be a talent for

collaboration. Thompson Owens, the Inner City support teacher who specialized in unit design and assessment, had encountered Ruth Sutton's work in Ontario and alerted his colleagues that she might meet these requirements. Further research into Ruth's work led us to invite her to meet the superintendent and school leaders, with a view to involvement in the on-going work. By the start of the 2000 school year, it had been agreed that Ruth would visit the district three times a year, for three weeks each time, for three years. Thompson would continue his support year-round. Feedback for Learning was chosen as the title of the work, to clarify its main focus.

The scale of this project was new to all of us. Some teachers thrived on the range of choices we offered, and others needed and welcomed a more structured approach. Towards the end of the first year, we developed a clear set of guidelines, scaffolding, which would support their learning until they no longer needed it. In the series *Creating Independent Student Learners: Practical Guides to Assessment for Learning*, Clarke, Owens, and Sutton (2006) detail these guidelines and lay out how the teachers' approaches to assessment and feedback developed in their classrooms.

At the end of each year of the project, we gathered information from a range of sources about the work and its impact. This allowed us to adjust our practice, and also encouraged us by indicating the positive responses of both students and teachers. In the second year, a further four schools joined in the work, and the number of teachers rose to 80. The natural movement of teachers into and out of these schools was a challenge. One benefit of this movement was that teachers committed to the work could share their learning with new colleagues in new schools, but the disadvantage was that we had to continually induct new teachers into Feedback For Learning to ensure that the work continued. We learned that we could not rely on newly-trained teachers knowing even the basics of effective assessment and feedback practice. This remains a frustrating concern.

In the third year, no new schools were invited to join the project, so as not to stretch support services too thinly, and Ruth devoted most of her time to support the school leaders while Thompson continued to work closely with the teachers. We knew that by the end of that year the schools, individually and collectively, should be able and willing to sustain their classroom and feedback strategies without extra support. To complement the accounts of classroom changes and their implications for teachers, we published a booklet for school leaders on leading a sustainable approach to learning improvement (Clarke, Owens, & Sutton, 2006).

Thinking and Working Smarter

Throughout Feedback for Learning we have helped our teachers to think and work "smarter," structure learning experiences that fully engage the learner, and, most of all, provide the steps for feedback. We have also paid attention to the steps school leaders need to take to recognize and support excellent

classroom assessment practice. Learning involves taking risks, supporting each other, looking for evidence of progress and adjusting their plans: Feedback For Learning is a multi-layered process influencing the behaviours of all the learners in the school, adult and child.

We now have all 21 Inner City District Schools involved. We monitor continually to see what works and what doesn't, asking what we should change, and what should remain constant.

We have published our work on Feedback for Learning, including the voices of our children and staff, to show the impact of this project on our schools. We continue to be impressed by the ability of our children and young people to reflect on their learning and to take responsibility for their achievement.

We undertook this work because we chose to, not because we were obliged to by some provincial or other imperative. The motivation was intrinsic, not extrinsic. Sustainability was the goal. Six years on, the strategies of Feedback for Learning are becoming the norm—the way our schools do business.

References

- Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan* 80(2).
- Clarke, P., Owens, T., & Sutton, R. (2006). *Creating independent student learners: Practical guides to assessment for learning—Nursery to 3; Nursery 4-6; Nursery 7-9; and School Leaders*. Winnipeg, MB: Portage and Main Press.



RUTH SUTTON has 10 years experience in classroom teaching and school management, and 20 years of continuous involvement with schools and school systems in the U.K. and across the English-speaking world. She consults in the areas

of curriculum management, assessment for learning, teacher development, school/college self review, and leadership training. Ruth writes and publishes books for teachers through Ruth Sutton Publications.



PAULINE CLARKE has worked as a teacher, school administrator, and superintendent. She has been the Superintendent of Schools—Inner City District, Winnipeg School Division, since 1990 where her priority is to build and sustain the work of

the group of Inner City Schools. Her most recent work emphasizes assessment for learning, integration of the arts, Aboriginal education, and building school leadership capacity. Ms. Clarke is a member of the United Kingdom Advisory Council on Learning Through the Arts.

Collaborating with Students in the Assessment Process

■ Damian Cooper

DAMIAN COOPER puts students at the centre of the assessment process and describes how teachers play the role of coaches to help students become self assessors who become expert at understanding their own learning.

Assessment is not something that is done to students; it is a process of collaborative communication in which information about learning flows between teacher and student. This two-way exchange of information is at the heart of assessment for learning. If assessment for learning is to improve the quality of students' work, then students must be involved in their own assessment.

If you have embraced this idea, you may have heard parents remark, "Surely assessment is the job of the teacher, not the students."

Such comments are sometimes made when parents are upset about what they perceive to be unfair grading of their child's work by a peer, or because they do not understand the purpose of peer assessment. That leads to an important guideline to follow when including students in the assessment process:

Train students in the skills of self and peer anecdotal assessment. Provide them with frequent opportunities to practise these skills to help them learn deeper and better. Do not require students to assign marks, either to their own work or to the work of their peers. Marking is part of evaluation (judgment) of student work and it is the responsibility of the teacher.

How Students Benefit

Students benefit from self assessment and peer assessment when they are participants in the process:

1. They come to understand the standards for quality work when they have to apply these standards themselves.
2. They learn to be less dependent on the teacher for feedback about the quality of their own work and, as result, become better able to independently monitor the quality of their own work.
3. They develop metacognitive skills so they are more able to adjust what they are doing to improve the quality of their work.

4. They develop and refine their capacity for critical thinking.
5. They broaden their own learning when they see how their peers approach a given task.
6. They practise and hone their communication and social skills when they are required to provide useful and meaningful feedback to others.

Teachers frequently ask, "How often should my students be involved in self and peer assessment?" The answer to this question is more a matter of purpose than frequency. What is the purpose of asking students to assess their own work, or the work of their peers? Decisions about who should assess a given task—teacher, student self assessors, peers, or a combination of these people—are best made with that question in mind. That said, there are certain points in the teaching/learning process when peer and self assessment are most effective:

- Include peer assessment as a routine practice when students are working on early drafts of written material. Peer assessment can be extremely effective in improving the quality of work before it is viewed by the teacher, provided that students have been taught what to look for in their partner's work, they have an assessment tool such as a checklist to focus their comments, and are required only to provide anecdotal comments and not scores.

The Eight Big Ideas of Assessment

1. Assessment serves different purposes at different times.
2. Assessment must be planned and purposeful.
3. Assessment must be balanced and flexible.
4. Assessment and instruction are inseparable.
5. For assessment to be helpful to students, it must inform them in words, not just numerical scores or letter grades.
6. Assessment is a collaborative process.
7. Performance standards are an essential component of effective assessment.
8. Grading and reporting student achievement is a caring, sensitive process.

—Damian Cooper

- Include self assessment as a routine practice in subjects like mathematics where students typically complete practice questions following the teaching of a new concept or procedure. Let students check their answers against the answers in the teacher's guide to determine whether they've "got it" or whether they need to review the new learning and complete more practice work. Scores from these practice exercises need not be recorded by the teacher—the purpose of such work is just *practice*. At a later date, there may be a test or assignment—the *game*—that will be scored by the teacher and will count towards students' grades.

Collaboration in Action: Two Examples

The following classroom examples illustrate how two teachers use a combination of teacher, self, and peer assessment to accomplish different assessment purposes. Sharon's language arts classroom incorporates self, peer, and teacher assessment at different stages of a writing task to improve students' work. In Dave's mathematics class, he has moved from an entirely teacher-driven assessment approach to one that emphasizes collaborative communication.

CASE STUDY 1 SHARON'S LANGUAGE ARTS CLASS

For some time, Sharon had been feeling that she was doing too much marking of students' written work. Every time she assigned them a task, however small, she felt compelled to collect it and take it home to mark. She raised this concern with the literacy consultant in her district. As a result of their discussion, she began to share the assessment responsibilities with her students.

Today, Sharon is teaching a lesson on improving the sensory impact of writing. She begins by showing students pairs of sentences and asking how the second sentence in each pair is better than the first. Here is one such pair:

Sentence Type 1

Marco walked up to the house.

Sentence Type 2

His heart pounding with each step, Marco inched towards the cobweb-covered door of the crumbling mansion.

Students work with their writing partners to discuss what makes the second sentence in each pair more effective than the first. After a few minutes, students share their observations about the sentences. Based on their responses, Sharon writes on the board a set of guidelines about writing descriptive sentences that appeal to the senses:

WRITING EFFECTIVE DESCRIPTIVE SENTENCES

Vary the syntax: don't always begin with the subject—e.g., "His heart pounding, Marco ..."

Vary the sentence length—e.g., short sentences create suspense.

Students as Self Assessors

Teachers play a key role in coaching their students to be self assessors. As self assessors, students need to:

- Know and understand how their teacher defines quality work.
- Know what the expectations are with respect to quality work and performance.
- Be trained in the use of assessment tools such as checklists and rubrics.
- Be coached in assessment process skills such as conferencing, self assessment, and peer assessment.
- Be reminded constantly, especially as they progress towards intermediate grades, that the purpose of completing work is to improve their learning, not to accumulate as many gold stars or points in as short a time as possible.

Use, but don't overuse, modifiers to provide sensory details—e.g., "crumbling."

Use concrete, specific nouns instead of generic ones—e.g., "mansion."

Next, Sharon provides students with five Type 1 sentences. Students read the sentences, and working alone, rewrite them as Type 2 sentences. After 15 minutes, there are opportunities for self, peer, and teacher assessments as students are asked to:

Step 1 Read the five sentences carefully and critically to ensure they are correct and that they meet the requirements listed on the board.

Step 2 Hand the five sentences to a writing partner and ask for feedback by indicating on each sentence:

- What makes it effective in appealing to one or more senses?
- What weakens its effectiveness?
- Give one suggestion as to how the sentence might be improved.

Step 3 Students conduct a brief, one-on-one writing conference during which they explain their comments to the writer. Each partner takes a turn providing this feedback.

Step 4 Sharon monitors students' progress on this peer assessment task, and after a reasonable length of time, students revise their five sentences, incorporating the feedback they received from their writing partners. The result of this activity will be polished sentences that are submitted for teacher assessment.

Can Very Young Children Self and Peer Assess?

Recently, I was asked by a Grade 2 teacher if I thought that five- and six-year-olds could self and peer assess reliably. I replied with a question, "Do you think it's important for young children to be engaged in self and peer assessment?"

"Absolutely!" she said without hesitation.

"So perhaps," I continued, "we should be less concerned with reliability at these early grades than we are with providing children with lots of opportunities to practice self and peer assessment."

"Yes," she said, "I understand what you're saying. Reliability is about consistency. But before we can expect children to demonstrate consistency when they are self or peer assessing, they need lots of practice."

I smiled. "Exactly. Reliability is of vital importance when we are talking about assessment of learning—in other words, assessment that is going to be used for reporting student achievement. But self and peer assessment by students in all grades should be limited to assessment *for* learning—assessment that helps them improve their work."

"OK, here's another question," she said. "What form should this assessment take, especially with primary students?"

I replied, "Descriptive feedback, such as 'This is what you did well,' and 'This is what you could do to make it better.' Children should not be assigning marks and numerical scores to their own work or the work of their peers."

"Absolutely, that's my job," she said.

"They'll also learn more about assessment and become more critical thinkers if they have to provide descriptive feedback instead of a score."

—Damian Cooper

HOW IS SHARON'S NEW APPROACH EFFECTIVE?

In Step 1, students look critically at their own work, checking their sentences against the guidelines on the blackboard. This step is essential, but often overlooked. Too often, teachers assess work students have not even read over and then must spend an enormous amount of time correcting careless errors.

In Step 2, students engage in thought-provoking peer assessment. They must critically examine the quality of their partner's work and provide useful and meaningful feedback. Notice: there is no requirement that students score each other's work. Numerical scores are not relevant to this process.

In Step 3, students practise both interpersonal and communications skills as they provide partners with feedback about their sentences. Sometimes peers disagree or one peer will want more feedback. This part of the process can help both partners clarify their thinking and understanding.

In Step 4, the teacher collects students' polished work to conduct her own assessment. Consider how much better the quality of this work will be compared to the work she might have received had she collected it without using the step-by-step process outlined here. Depending on her purpose for the assessment, and at what point in the term this takes place, the teacher's assessment may be purely anecdotal (similar to the student assessment), or it may include a letter grade or level.

CASE STUDY 2 DAVE'S MATHEMATICS CLASS

Dave had been teaching mathematics in a consistent, and somewhat traditional way, in junior grades for 10 years. Dave believed that assessment meant measuring what his students had learned, and it was entirely the responsibility of the teacher. In recent years, Dave has become increasingly frustrated and disappointed with his students' poor achievement in math. These are his concerns:

- Despite his best teaching efforts, many students do not understand the math concepts.
- Although he assigns homework on a daily basis, many students fail to complete it.
- He is constantly overwhelmed by the marking that he feels compelled to do, yet he knows it is having little impact.
- He is invariably disappointed by the poor performance of many students on the bi-weekly tests.

Dave visited a school known for teaching excellence in mathematics. Dave spent the day observing one teacher and her students and chatting with her about her success. As a result of this visit and subsequent communication with his colleague, Dave implemented a number of new strategies in his class.

HOW IS DAVE'S NEW APPROACH EFFECTIVE?

Dave now constantly checks to see whether the whole class, as well as individual students, are "with him" when he is teaching a new concept or procedure. Checks for understanding include:

orbit PROFESSIONAL DEVELOPMENT

MAKING CLASSROOM ASSESSMENT WORK FOR STUDENT LEARNING

Carol Rolheiser and Barb Bower

The professional development activities outlined in this pullout are designed to foster leadership and learning focused on classroom assessment that supports student learning. The activities can be facilitated in courses, workshops, or job-embedded professional learning experiences. The suggested processes that focus on the articles from this issue of *Orbit* will foster important discussion and spark key questions. Hopefully, dialogue among readers will result in the generation of positive actions for classrooms, schools, districts, and universities.

■ Learning Activity #1

3-2-1 Graphic Organizer and Walk-About

Led by: Teacher Leaders/Principals/Vice-Principals

Topic: Assessment for Learning: Why, what, and how?

Participants: Principals/Vice-Principals/Teachers

Time: 60 minutes

Preparation:

For each participant, provide a copy of the article *Assessment for Learning: Why, What, and How?* by Dylan Wiliam (other articles in this *Orbit* issue can also be used for this learning activity), and the 3-2-1 Graphic Organizer.

Instructions:

- Participants are provided with the 3-2-1 Graphic Organizer handout and invited to record notes on it as they read the article.
- When participants have finished reading and recording their ideas, they are invited to do a Walk-About. When signalled, each person stands and connects with one other in the room, sharing three key ideas from the article. A signal is provided and each person finds a second partner, proceeding to share the two key questions that emerged from reading the article. A final signal is provided and participants find a third partner, sharing one possibility for his/her own practice based on the article.
- Large group debriefing: What were some of the key ideas from the article that you or your partners shared? Why were these ideas important? What key questions did you discuss? What possibilities emerged as you thought about your own practice and context?

3-2-1 Graphic Organizer

3 Key Ideas

-
-
-

2 Questions

-
-

1 Possibility for my Practice

-



■ Learning Activity #2

4-Session Study Group

Led by: Teacher Leaders/Principals/Vice Principals/Superintendents

Topics:

1. Assessment and Learning
2. Assessment for Learning in Action
3. Rethinking Classroom Assessment

Participants: Teachers or principals

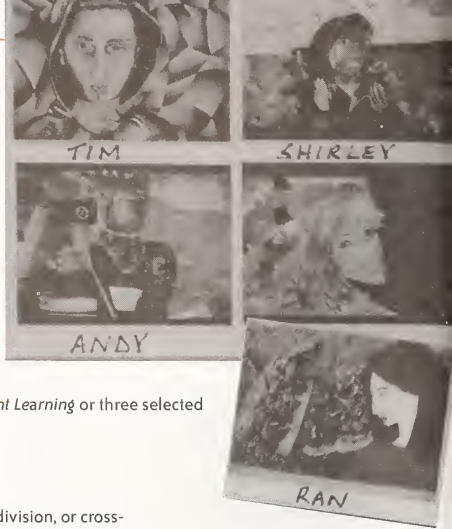
Time: 4 sessions of 45–60 minutes each

Preparation:

1. Provide copies of the *Orbit* issue *Making Classroom Assessment Work for Student Learning* or three selected articles from the issue, one from each of the following sections:
 - a. Assessment and Learning
 - b. Assessment for Learning in Action
 - c. Rethinking Classroom Assessment
2. Form study groups of six participants (teacher participants can be grouped by division, or cross-grade/division/subject areas).
3. Chart paper and markers will be required for the fourth week.

Instructions:

1. Share the purpose for using a 4-Session Study Group with participants, e.g., "These sessions will provide an opportunity for professional dialogue focused on the topic of *Making Classroom Assessment Work for Student Learning*." Share the following desired outcomes:
 - Engage in professional dialogue focused on professional reading.
 - Increase understanding and help bridge the gap between theory and practice.
 - Practise one strategy for effective group discussion.
 - Practise taking different roles in group discussions.
 - Consider applications and next steps.
2. Form study groups of no more than six people per group. Each group reviews the following six procedures:
 - Read and understand the specifics of each role.
 - Choose one or two roles for each week—different roles for each session.
 - For each session, read the assigned article in advance.
 - Come prepared for assigned role(s) for each session.
 - After each session, talk with colleagues in the study group about the article and generate new understandings.
 - In Week 4, the study group works together to create a visual as a summary or synthesis of what has been learned. Visuals can be shared with a larger group (e.g., other school staff, other study groups).
 - Debrief the study group process. Participants can reflect on the following questions:
 - What are some positive things you noticed/learned from your involvement in the study group sessions?
 - How did you feel about your participation in this learning opportunity?



What insights do you have as a result of this learning opportunity?

What suggestions do you have to assist in planning for future study groups?

3. The study group sessions are scheduled for approximately one hour, four times during a given block of time (e.g., at a lunch hour, after school, early morning for a school staff OR at four professional development sessions scheduled by a District leader). Times are negotiated.

Select one article from Section 1 of this *Orbit* issue. Assign roles using the following outline:

□ WEEK 1 Read: _____

Meeting Date/Time: _____

Roles: (*Notes mandatory roles for each study group)

- A *Discussion Facilitator
- B *Recorder
- C *Meeting Manager
- D Summarizer
- E Passage Master
- F Connector

□ WEEK 2 Read: _____

Meeting Date/Time: _____

Roles: (*Notes mandatory roles for each study group)

- A *Discussion Facilitator
- B *Recorder
- C *Meeting Manager
- D Summarizer
- E Passage Master
- F Connector

WEEK 3 Read: _____

Meeting Date/Time: _____

Roles: (*Notes mandatory roles for each study group)

A *Discussion Facilitator:

B *Recorder

C *Meeting Manager

D Summarizer

E Passage Master

F Connector

WEEK 4 Design a visual summary

For this week, the study group comes together to create a visual (e.g., mind-map, fishbone, concept map) as a summary or synthesis of what has been learned. The visuals and group debriefing may also be shared with others.

STUDY GROUP ROLES

A *Discussion Facilitator

Your task is to help people talk about the "big ideas" in the readings and to help them share their reactions. Call on participants in their different roles to share what they have prepared. In order to facilitate the discussion effectively, you might prepare a set of questions that you believe would help your group discuss the reading, or you might use some of the questions below to generate discussion.

Your main job is to encourage, verify, assure equitable participation, and effect closure to the meeting (*keep your eye on the clock*).

Sample questions (Schwartz & Bone, 1995):

Retell:

- What was discussed here?
- Can someone summarize briefly?
- What are the one or two most important ideas, issues, concerns?

Relate:

- What was going through your mind while you read this?
- How did you feel while reading this?
- Did today's reading remind you of any real-life experiences in your classroom/school?

Reflect:

- What questions did you have when you finished this article?
- Did anything in this article surprise you?
- Predict some things you think will be talked about next.
- What did you learn from this reading? Any ah-ha's (great ideas)?

B *Recorder

Your job is to act as recorder for the study group meeting, taking notes on important points and ideas discussed. Point form notes and headings are acceptable. This will serve as a record of your meeting. At the beginning of each subsequent session, please be prepared to distribute your notes from the previous meeting to all members. You may wish to begin the next meeting by reviewing the "minutes."

C *Meeting Manager

Your job is to organize housekeeping details for the study group meetings. You need to clarify the time and location of the sessions and bring any materials that the group will need. For example, for Week 4, you are responsible for bringing the materials you will need to create the visual for your group. You should also think about providing a small treat, coffee, etc., to create a comfortable atmosphere.

D Summarizer

Your job is to prepare, in advance, a brief summary of the main sections of the text. The other members of your group will be counting on you to give quick statements that convey the key points, main highlights, and the essence of the session's reading. If there are several main ideas or events to remember, you may wish to create a one-page handout containing a list, web or chart that highlights the main points. (*If you do prepare something, please give to the Recorder to include with the minutes.*)

E Passage Master

Your job is to identify two to three special sections in the article that warrant additional discussion. You decide in advance which passages or paragraphs are worth reviewing and then think about how they could be shared with the group. You can read passages aloud yourself, ask someone to read them, or have people read them silently and discuss afterwards.

Possible reasons for choosing a passage to be shared is that it is:

- Important
- Surprising
- Confusing
- Informative
- Controversial
- Thought-provoking
- Other?

1. Location: Page: _____ Paragraph: _____

Reason(s) for Choosing: _____

2. Location: Page: _____ Paragraph: _____

Reason(s) for Choosing: _____



F Connector

Your job is to explore two to three connections between the material you are reading and your experiences. This means connecting specific parts of the reading to similar events, your classroom or work environment, your understandings, etc. You might also see connections between this material and other writings on a similar topic, or by the same author. Whatever your connections ... they are worth sharing!

1. Location: Page: _____ Paragraph: _____

Reason(s) for Choosing: _____

2. Location: Page: _____ Paragraph: _____

Reason(s) for Choosing: _____

■ Learning Activity #3

Critical Thinking Forum

Led by: Principals/Teacher Leaders

Topic: Any topic selected from the range of articles in the *Orbit* issue, *Making Classroom Assessment Work for Student Learning*.

Participants: Teachers

Time: 45 minutes

Preparation: Provide a copy of one article per participant.

Instructions:

1. Arrange participants in groups of three.
2. Choose one organizer to use from the following samples: A, B, or C.
3. Provide participants with the article.
4. Provide participants with the selected organizer for recording their ideas.
5. After individual reading and recording, each person shares ideas Round-Robin style, focusing on one section of his/her critical thinking organizer. Continue Round-Robin until all sections are debriefed.
6. At the end, discuss similarities and differences of ideas among all three group members.

7. **OPTIONAL:** The facilitator can highlight the importance of a professional dialogue such as this by posing the following questions:

- Why do we need to be critical reviewers of professional reading?
- What have we learned from this process and from each other?
- How might this process be adapted for encouraging critical thinking by students?

Organizer A: POP!

POP!

I am:

Passionate about...

Opposed to...

Perplexed about...

Organizer B: Agree/Disagree Chart

Idea I agree with...

Idea I disagree with...

Idea I want to think more about...

Organizer C: Thinking Stems

An idea that matches with my own thinking is...because...

An idea I want to consider further is...because...

An idea that causes me concern is...because...

A question I would like to pose to the author(s) is...because...

Idea I disagree with...

Idea I want to think more about...

Reference

Schwartz, S., & Bone, M. (1995). *Retelling, relating, reflecting: Beyond the 3 R's*. Concord, ON: Irwin.



Carol Rolheiser is Associate Dean, Teacher Education, at OISE/UT. **Barb Bower** is Lead Mentor at The Bishop Strachan School in Toronto, ON. Carol and Barb have collaborated professionally since 1992. Along with conducting joint research on student self-evaluation and school/system capacity-building, they co-authored the ASCD publication, *The Portfolio Organizer: Succeeding with Portfolios in Your Classroom* (2000).



- **Show me your thumbs**—Dave stops part way through the lesson and says, "If you understand completely, thumbs up; if you understand partially, thumbs straight across; if you don't understand, thumbs down. Now, show me your thumbs." Dave scans the room and notes which students have thumbs down, or thumbs horizontal. He instructs those who understand to move ahead, and then groups the other students in a corner of the classroom and provides further instruction targeted at their misunderstanding.
- **One-on-one questioning**—During whole-class instruction, it is hard to know which students are not following the lesson. Many students who have difficulty understanding a concept are reluctant to raise their hands to ask questions. Once Dave has completed a lesson on a concept and students are working, he moves from desk to desk, speaking individually to students, especially those who have been having trouble with this topic, and asking questions to check for understanding.
- **Conferencing**—As a regular part of his classroom routine, Dave conducts brief, one-on-one conferences with his students. While these conferences serve many purposes, a key purpose is to conduct periodic checks on each student's understanding. His students know when conferencing days fall, and they know what Dave's behavioural expectations are for the rest of the class during these sessions.

SELF AND PEER ASSESSMENT

Self checking—After a verbal check for understanding a new concept or procedure, he assigns practice questions from the text. Students then check their answers against those in the back of the teacher's guide. If they are mostly correct, they move on. If they get several wrong, they work with a partner to try and improve their understanding.

Consolidating understanding—Dave now builds into every lesson the opportunity for students to talk through their understanding with their peers. While this does result in a noisy classroom, he knows the difference between "good noise" (the chatter of students sharing their learning) and "bad noise" (off-task chatter).

As Dave compared his past practice to his new approach, he realized that his focus had been teaching curriculum rather than teaching students. He had been driven more by pressure to "cover the outcomes" than by a commitment to learning for all. If the purpose of teaching is to "uncover understanding" (Wiggins, 1998), then Dave knew he had left that purpose behind.

Although these are recent changes in classroom routines, Dave has already noticed a big improvement in his students' attitudes, especially in those who had been struggling. No longer does he hear, "I just don't get it! I'm useless in math!" Long before a test, students know that their teacher is going to check their level of understanding and address any misconceptions they may have.

Summary

These two case studies illustrate the power of collaboration in the assessment process. Sharon realized that if she involved her students in self and peer assessment, she could save herself a lot of marking, while at the same time, significantly improve the quality of the work that she did mark. Dave discovered that by introducing a feedback loop into his mathematics class, he was able to monitor his students' understanding of concepts and to respond immediately to those who were having difficulty, rather than waiting for poor test results to indicate this. By introducing self and peer assessment into his daily routines, he deepened mathematical understanding and lessened his marking load.

This article has explored the benefits of involving students directly in their own assessment as well as assessment of their peers:

- Self and peer assessment help students to internalize the standards for quality work.
- Through being involved in the assessment process, students develop the capacity to be objective about the quality of their own work.
- When students are taught how to self and peer assess, the quality of work that they submit for marking improves.
- Effective and routine use of self and peer assessment can significantly reduce the amount of marking for teachers.
- Very young children can learn how to self and peer assess.

However, teachers should be mindful that self and peer assessment are assessment for learning strategies and, as such, they need to take the form of descriptive feedback.

The evaluation and grading of student achievement for reporting purposes is the responsibility of the teacher.

Note

This article is adapted from Cooper, Damian (2006) *Talk about assessment: Strategies and tools to improve learning*. Toronto: Thomson, Nelson.

References

- Black, P. & William, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139–144; 146–148.
Wiggins, G., & McTighe, J. (1998). *Understanding by design*. Alexandria, VA: ASCD.



DAMIAN COOPER is an independent education consultant who helps schools and school districts improve their instructional and assessment skills. He has taught in academic environments, vocational settings, and schools serving students with severe learning disabilities. Damian has also worked on ministry and consortium development teams, consulting on and contributing to a variety of curriculum and assessment initiatives.

Dilemmas of Policy, Research, and Practice

■ John Myers

JOHN MYERS explores the knotty issue of group work and how assessment fits, with particular attention to the dilemmas of grading group work.

In the area of grading students working in groups there are always tensions among:

- what policy makers suggest or mandate;
- how such mandates are interpreted by teachers, schools, and school districts; and
- how teachers and policy makers pay attention to or ignore what research says.

These can be represented thus.

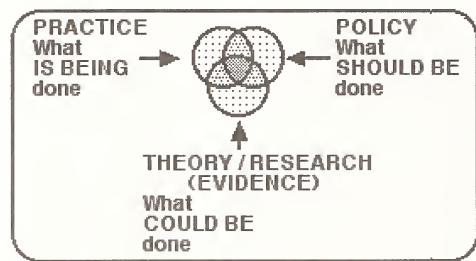


Figure 1

The model is dynamic in that these borders are always shifting with changes in practice, shifts in policy, and new evidence from research. The challenge for all of us is for better alignment among these components.

We never get full alignment.

The tensions appear in most areas of teaching, curriculum, and assessment. This article looks at the interplay in the area of grading students working in groups.*

Some argue that grading should always be individual on grounds of fairness (O'Connor, 2002; Kagan, 1995). Others (Johnson & Johnson, 2004; Ledlow, 1999) argue that group

grading is not only fair but also when done well can increase student achievement. Baloch (1998) and Vermette (1998) suggest that group grading may be appropriate under certain circumstances.

Ministry policy (see <http://www.edu.gov.on.ca/eng/curriculum>) is silent on this issue though interpretations of the policy are often opposed to group grading (see Toronto District School Board, 2006; Kawartha Pine Ridge DSB, 1999). In any case, teachers will use some form of group grading despite what the "experts" might say because it is easier to grade six products than 30!

The following two options for group grading represent defensible options. Neither option is perfect.

Option 1: Individual Products Involving Group Practice or Group Study

In a typical case, students study as a group but take individual tests or hand in individual essays. It seems to be the "fairest" system for an individual student and represents what teachers traditionally do. Yet it may not induce co-operation or teamwork among members of a group. If I must take an individual quiz and I am particularly good at taking individual quizzes, why should I study with others if I don't need to?

HOW CAN WE KEEP INDIVIDUAL GRADING IN THESE SITUATIONS, YET ENCOURAGE CO-OPERATION WITHIN GROUPS?

One strategy uses bonus scoring in which all members of the team earn bonus points or public team recognition for achieving a minimum team average or a minimum individual score. This is akin to a "playoff bonus" in a professional team sport in which each member of the championship team gets the same reward though individual contracts may vary greatly (Vermette, 1998). Another version of this system rewards perfect scores or improved scores, in which students' quiz scores are compared to past averages. Points are awarded based on improvement over previous scores. These points are totaled to form team scores. Teams meeting certain criteria such as a minimum average gain certificates, bonus points, or other forms of public recognition (Slavin, 1986).

This form of reward interdependence seems to result in significantly greater achievement by struggling students of low

Conditions for Effective Group Assessment

The following conditions are needed for grading group-based work:

- Specific links to learning outcomes are stated clearly to students.
- Students have input into outcomes, understand the criteria, and recognize fairness.
- Accountability is built into the work of individuals and groups.
- Minimum criteria for co-operatively-structured groups are set—i.e., the task is suitable for group work and structures positive interdependence, individual accountability, and face-to-face interaction.
- Students are given an opportunity to reflect on learning.
- The teacher provides supportive, specific, and immediate feedback related to ongoing formative assessment of process and product.
- There is a positive classroom climate and sufficient experience by students in conducting formative assessment, including self and peer assessment.

—John Myers, 2003

socio-economic backgrounds when tested on their understanding of basic knowledge (Slavin, 1995). Nevertheless this system has its critics (Kohn, 1994).

Another approach to bringing individual accountability to group work is the use of random marking or numbered heads in which individuals in a group complete the product and the teacher marks one product from each group; that mark is awarded to every member in the group.

The Johnsons (2002) have additional suggestions, including opportunities for students to take tests as a group as a feedback and improvement measure after individual tests are taken.

Some of these options are more palatable to teachers, students, or parents than others. Remember the preconditions for any group assessment. My advice is to keep the grading individual if the product is individual. Use the power of co-operative groups to help all students achieve when they practise

a skill or procedure, when they explore an idea, or when they bring different perspectives to the table.

Option 2: Grading Group Products

The options above apply to traditional testing. What can we do with alternative assessments involving performance tasks or “authentic” assessment? These tasks often result in group-based products. If “authentic assessment goes hand-in-hand with co-operative learning” (Vermette, 1998), then how do we grade the work based on the following example?

FACT: Canada makes a lot of maple syrup: about 75% of the world's total. In 1994 we made 19,500 kilolitres of the gooey stuff worth about \$95 million. About 80% of our syrup is produced in Quebec.

You are members of an advertising agency which has just been given a contract to develop commercials for the 2010 World Cup soccer final in South Africa. This is an event which will be seen by more than a billion people around the world, the most-watched event in television history.

Prepare a 30-second television commercial to advertise Canadian maple syrup to the world and persuade people to buy it. The audience for the commercial may be from any country or region, and speak in any language, since over 100 television networks will carry the final game to almost every nation in the world.

The task must meet the following conditions:

1. The maple syrup facts must appear in your commercial in some form.
2. You must find a way to translate these large numbers into an image that viewers can actually understand. You may do this verbally, through a picture or diagram, through drama, song, etc. Your goal is to make the large numbers really mean something.
3. You may ask others in the room for help. You can leave the room and use a phone or any resource you find helpful.
4. You may use calculators.
5. You can round off numbers or use estimates that you believe to be very close to the real thing.
6. You cannot ask the teacher for help.
7. You have 40 minutes to prepare your presentation.
8. Your presentation cannot last longer than 30 seconds.
9. You must involve every member of your group in the presentation.

In the above case you may need to reward both individual and group achievement. One advantage of doing so is to have a more complete assessment of an individual's achievements. Another advantage comes from the positive interdependence promoted to accompany individual accountability. One disadvantage is that such combined grading is not “traditional” and thus may take more work on the part of teachers and students so that the necessary preconditions are met.

Here are some grading possibilities, each of which has its merits and its weaknesses:

- A performance ratio is established whereby two-thirds of the marks are based on individual effort and one-third on the average performance of group members.
- The teacher awards a mark for each group member for a group product, with group members contributing an additional mark based on self and/or peer evaluation of their individual efforts. Individual contributions must be discernible. In the example above, every member of the group participates in making the commercial, and in the group evaluation, the work members do in the preparation of the commercial can be identified by all of the evaluators.
- The teacher awards a mark to the group who then distributes the mark among its members according to the individual contributions made. In this scenario, if a group of three gets a mark of 16/20 or a total of 48 marks out of 60, they discuss the distribution of the marks and hand in a signed form stating the individual marks and the justifications behind the award.

It is important, in addition to all of the preconditions, to build in an additional safeguard or two. Here are five safeguards to use with the sample task above:

1. Make sure the group for this task is no larger than four. The larger the group, the easier it is for students to hide.
2. Be careful about how the groups are formed. Vermette (1998) has a detailed examination of this important consideration.
3. Monitor the groups and make notes on group behaviour. In particular, observe the groups you feel are "struggling."
4. In the case of oral presentations, have the class as the audience rate the groups, usually based on a marking scheme or simple rubric for evaluating communication and presentation. Look for the accuracy of the content, in this case, the math.
5. ALWAYS reserve the right to renegotiate—not "change"—the internal group evaluation if the self evaluation and the evaluation of the rest of the class, including the teacher's evaluation, are at serious odds. This happens very rarely and, in my own teaching career, in such instances the group has convinced me that their grading is justified.

So What Do I Do Now?

The short answer is to use your best judgment, based on what you know about effective group work, quality assessment, and evaluation to work through the dilemmas and work with the options as best you can.

In addition to the "playoff bonus" analogy above, consider the following: Most grades in K–12 are formative since most students don't really get assessed for sorting purposes except for senior high (Grade 12), and possibly Grades 8–9. So use all the assessment data you can to promote growth. And in these "non-sorting" grades, Anne Davies (2006) suggests that these components of evidence could all be used in the grading process:

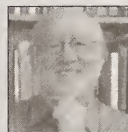
- products,
- observation of process, and
- conversations, listening, and reading.

Note

*If the groups are not effective in the first place, assessment will be flawed (Dippong, 1992). See www.glacie.ca for an introduction and links to the literature on effective groups.

References

- Baloché, L.A. (1998). *The co-operative classroom: Empowering learning*. Upper Saddle River, NJ: Prentice Hall.
- Davies, A. (2006). Making classroom assessment work for learning. Handout for Assessment, Evaluation, and Reporting Across the IB Continuum (Nov.) Toronto.
- Dippong, J. (1992). Two large questions: Teacher challenges and appropriate student tasks in assessing and evaluating student learning. *Co-operative Learning*, 13(1), 6–8.
- Johnson, D.W. & Johnson, R.T. (2004). *Assessing students in groups: Promoting group responsibility and individual accountability*. Thousand Oaks, CA: Corwin.
- Johnson, D.W. & Johnson, R.T. (2002). *Meaningful assessment: A manageable and co-operative process*. Boston: Allyn & Bacon.
- Kagan, S. (1995). Group grades miss the mark. *Educational Leadership*, 51(8), 68–71.
- Kawartha Pine Ridge District School Board. (1999). *A resource for assessment, evaluation, and reporting*. Peterborough, ON: Kawartha Pine Ridge District School Board.
- Kohn, A. (1994). *No contest: The case against competition*. Boston: Houghton-Mifflin.
- Ledlow, S. (1999). *Group grades in co-operative learning classes*, Arizona State University. Accessed May 19, 2006 from <http://www.public.asu.edu/~ledlow/sledlow/group.htm>
- Myers, J. (2003). Assessing citizenship and character using co-operative learning. *Orbit* 33(3), 47–49.
- O'Connor, K. (2002). *How to grade for learning: Linking grades to standards*, (2nd ed.). Arlington Heights, Ill: Skylight.
- Slavin, R.E. (1986). *Using student team learning* (3rd ed.). Baltimore, MD: Center for Research on Elementary and Middle Schools, The Johns Hopkins University.
- Slavin, R.E. (1995). *Co-operative learning* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Toronto District School Board. (2006). *Fresh AER: Assessment, evaluation, and reporting in secondary schools*. Toronto: Toronto District School Board.
- Vermette, P.J. (1998). *Making co-operative learning work: Student teams in K–12 classrooms*. Upper Saddle River, NJ: Merrill.



JOHN MYERS is a curriculum instructor at OISE/UT. He has taught primary to adult education in three provinces and three countries over three decades. In keeping with the "3" theme, his teaching interests include: group processes in the classroom; assessment and evaluation; and helping teachers broaden and deepen their instructional repertoire.

ASSESSMENT TRACKING BOARDS:

A Focus for Professional Dialogue■ *Mary Lou McKinley, Kathy Witherow, and Angela Puma*

MARY LOU MCKINLEY and her colleagues show how teachers in Ontario schools are using tracking walls as a stimulus for professional discussion about each student's learning and what is required to move the learning forward.

Using professional learning communities to discuss student achievement is becoming a focus within many districts in Ontario. The purpose of posting "evidence of" student learning on a data wall is to promote discussion around the multifaceted types of information needed to analyze student learning. This, in turn, has the potential to become an effective feedback system for professional dialogue (Ross, 2004). Stakeholders within the professional learning community become collectively responsible for supporting all students to achieve based on a pre-determined focus which informs instructional practice.

Primary Reading Scores on Display

The majority of districts in Ontario are asking primary teachers to regularly submit data collected on student reading levels based on formal running record assessments. A running record is a formative literacy assessment tool. It provides information about reading behaviour during the reading process. Subsequently, many schools have developed a visual representation of student progress: assessment tracking boards (or data walls) that display the results of these assessments (see Figure 1).

Some schools are using assessment tracking boards as an integral component of focused discussion during their professional learning community meetings. Often tracking walls simply display individual student reading levels as determined by a formal reading assessment tool. Students who are not meeting the pre-determined targets and goals set by the division, school, and/or district are identified as being "at-risk." The discussion surrounding the evidence of student learning as displayed by the tracking board serves the purpose of starting the discussion and setting an action plan by:

- identifying students who appear to be "stuck" and/or struggling;
- identifying the strengths and gaps of students by collecting and examining other pertinent data;

- ensuring that assessment and instruction are aligned by identifying, discussing, and monitoring instructional approaches based on the analysis of the data and the needs of the students identified;
- determining and reassessing the teaching strategies chosen as determined by current best practices and the analysis of the data collected through ongoing observation of the students engaged in the reading process—reading, responding, and interacting with print; and
- monitoring and reflecting upon the progress of students by analyzing the complexities involved in the student's behaviours during the reading process by asking, "Is the student gaining self-confidence and competence by applying strategies in an independent manner?"

Although the data tracking wall displays the reading level attained, professional dialogue needs to look at the evidence of learning collected on repeated measurements of daily reading tasks that the child is undertaking. As Clay (1993) suggests, a running record is an observational tool to collect data on learners at work; the process (especially with emergent readers) is as important as the final product.

Information is gathered in a systematic manner and documents a child's behaviour during the reading process for:

- competencies and confusions;
- strengths and gaps; and
- evidence of a child's understandings.

Running records are the window into how a child interacts with text; teachers need to be able to understand and interpret

Professional Learning Communities

- Focus on learning
- Focus on results
- Collaborate (replacing teacher isolation)
- Use data for relevant information for decision making
- Share leadership

—Adapted from DuFour & Eaker, 2002

data collected through interactive sessions, observations, and samples of student performance, as these are key to providing effective and timely instruction for all students (Stiggins, 1991).

One School's Experience

One primary division in a small northern Ontario school uses a simple framework (see Figure 2) to plan instructional support for those students identified as being at risk and/or struggling in learning how to read. Students from Grades 1 through 3 are initially identified based on the student's reading level as determined by a formal running record posted on the assessment tracking wall. The tracking wall is a visual display of each student's reading level (see Figure 1). The background and the student records are colour co-ordinated which instantly provides a visual clue as to those children who are not within the identified and/or acceptable reading range as determined by the current research community (Fountas & Pinnell, 1999).

After the initial identification of at-risk students, an aggregate of data is collected and presented by classroom teachers during the initial professional learning community's meeting in September. By analyzing the data collected and then discussing, and prioritizing student needs, some students are designated to work within a booster group while others are supported through suggested changes within individual classroom settings.

This school supports identified students through a literacy team approach. The team comprises two special education resource teachers (one full-time and one part-time assignment) and two educational assistants. They provide ongoing, flexible support to students identified by the primary division during the professional learning community meetings—the team works as a unit admitting and demitting students based on ongoing

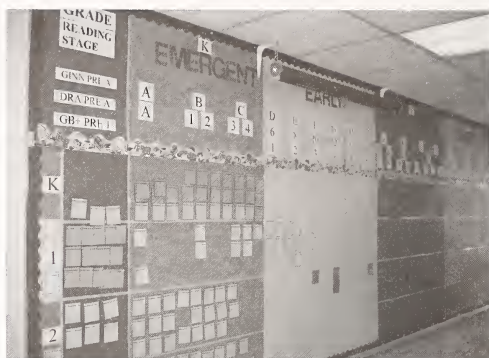


Figure 1: Visual Display of Running Records

This tracking board is from a dual track school in northern Ontario and is located in the primary division book room.

discussions and monitoring of student progress by the team, other teachers in the division, and the classroom teacher. The entire division shares the responsibility for ensuring that individual students receive the support best suited to augment their achievement in reading.

Ongoing observations and the collection of complementary assessment data by the classroom teacher and literacy support staff are an integral component of the monthly professional learning community meetings to monitor whether the student is starting to use and apply specific strategies in unique learning situations within the classroom setting. As the process continues,

Time Frame and Identification of Students	Reading Behaviour	Focus for Instruction	Who and How
- Students are identified as being at risk based on running records, targets, and identified goals as determined by the school and/or district. - Time frame to admit and demit within each grade level—monitored by an aggregate of assessment data and reviewed during bi-weekly professional learning community meetings.	- Identification of the strengths that the student brings to the reading experience and gaps identified as those that the student is not applying in authentic situations. - List of reading behaviours and expectations according to individual grade levels.	- Brainstorming and identifying the specific strategies to be employed based on the gaps and/or inconsistencies in application of strategies as noted through assessment(s) and/or observations of classroom teacher. - Explicitly identify strategies and resources needed to address the gaps. - Monitoring the effectiveness of programming based on formative assessments.	- Special education and/or resource teacher to work with individual and/or small groups of children as identified through process. - Specific strategies identified through group consensus during professional learning communities.

Assessment Tracking Wall

Tracking walls promote professional dialogue by:

- Identifying and setting goals for all students based on current research (Fountas & Pinnell, 1999).
- Identifying struggling students and planning next steps to address gaps in learning.
- Engaging stakeholders in professional dialogue on reading acquisition by:
 - examining instructional practices to support the identified gaps in student performance ... to be reassessed during subsequent professional learning community meetings (ongoing dialogue based on formative assessment),
 - monitoring strategies that are being applied during focused instruction with individual students.

evidence of student learning is the basis for discussion and professional dialogue among the primary division during professional learning community meetings (Schmocker, 2003; Shepard, 2005).

Aligning Assessment to Inform Practice

This discussion focused on using an assessment tracking wall to plan strategically for individual students as determined by ongoing formative assessment data and professional dialogue during professional learning community meetings. The monthly and/or bimonthly professional learning community meetings are pivotal for the primary staff and administration to explore questions critical to student success in reading achievement by explicitly articulating:

- indicators of success for the stages of reading development;
- individual student strengths and gaps as measured by the aggregate of data collected and analyzed; and
- instructional practices and supports to address the gaps identified—differentiated instructional approaches, personnel to deliver, monitor growth, and check for success.

Nevertheless, Ross (2004) cautions:

"... assessment is tightly bound to deep-seated teacher beliefs about evidence, student motivation, and instructional effectiveness." (p. 187)

"The most promising strategy for sustained, substantive school improvement is developing the ability of school personnel to function as professional learning communities."

—Dufour & Eaker, 1998, p. xi

and suggests that improvement in student learning will be dependent upon an entire division using a data assessment board as the lever for focused discussion regarding the delivery of curriculum at the classroom level.

References

- Clay, M. (1993). *An observation survey of early literacy achievement*. Portsmouth, NH: Heinemann.
- Dufour, R., & Eaker, R. (1998). *Professional learning communities at work: Best practices for enhancing student achievement*. Alexandria, VA: ASCD.
- Fountas, I.C., & Pinnell, G.S. (1999). *Matching books to readers ... Using leveled books in guided reading, K-3*. Portsmouth, NH: Heinemann.
- Ross, J.A. (2004). Effects of running records assessment on early literacy achievement: Results of a controlled experiment. *Journal of Educational Research*, 97(4), 186–194.
- Schmocker, M. (2003). First things first: Demystifying data analysis. *Educational Leadership*, 60(5), 22–25.
- Shepard, L.A. (2005). Linking formative assessment to scaffolding. *Educational Leadership*, 63(3), 70–75.
- Stiggins, R.J. (1991). Assessment literacy. *Phi Delta Kappan*, 72(7), 534–539.



MARY LOU MCKINLEY, PhD, is a Student Achievement Officer with the Literacy and Numeracy Secretariat, Ontario Ministry of Education.

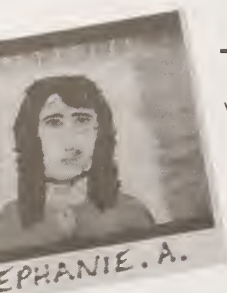


KATHY WITHEROW is currently a Principal of Leadership Development with the York Region District School Board. She returned to YRDSB after completing a recent secondment with the Turnaround Team program, Ontario Ministry of Education.



ANGELA PUMA has recently retired from the TCDSB as an Assessment and Programming Teacher. She is currently a Student Achievement Officer with The Literacy and Numeracy Secretariat, Ontario Ministry of Education.

Rethinking Classroom Assessment



This final set of articles demonstrates the value of serious reflection. The authors, all seasoned educators, have continued to sharpen their understanding of the meaning, value, and power of quality assessment. After careful thought, they share their insights, personal experiences, and knowledge of assessment theory and practice.

- **Patricia McAdie and Ruth Dawson** challenge the merit and worth of large-scale testing. They argue that these tests can cause students to experience unnecessary stress and teachers to feel pressured to have their students perform at a high level. They lament the fact that one test of learning is inappropriately used to evaluate a range of disparate goals. What we need, argue the authors, is a clarification of the goals of assessment and recognition of the importance of daily assessment for learning within professional learning communities.

Standardized Testing, Classroom Assessment, Teachers, and Teacher Unions

■ *Patricia McAdie and Ruth Dawson*

PATRICIA MCADIE and RUTH DAWSON argue that large-scale testing can cause students unnecessary stress and teachers to feel pressured to have their students perform at a high level.

Teacher unions across Canada oppose the use of standardized testing to evaluate students, teachers, and schools. Teacher unions take this position based on a careful reading of the research. There are countless studies detailing problems with the use of standardized testing, whether or not it is high-stakes (see, for example Froese-Germain, 1999; Kohn 2000; Amrein & Berliner 2002).

Teacher unions also represent the views and experiences of their members in opposing standardized testing (see sidebar pg. 32).

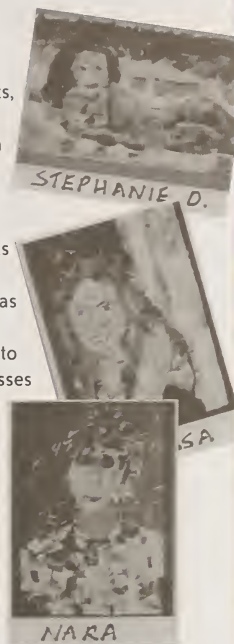
These tests impact students, schools, and communities. Many students experience unnecessary stress surrounding the testing, regardless of how teachers emphasize a relaxed approach. When students are not allowed to ask questions for clarification and teachers cannot provide guidance this is an unnatural environment, particularly for students in Grade 3. There is pressure to have students achieve at a high level on the tests in order to raise the school's results and their rank. While EQAO maintains that ranking schools based on the EQAO results is not appropriate, such ranking happens and often attracts much media attention.

- **Louis Volante** provides practical strategies and suggestions for promoting fair classroom assessment practices. Using reliability and validity as a framework for reducing bias, he develops a foundation that supports improved instruction and learning. These strategies for reducing bias not only respect the diverse needs and backgrounds of students, but they also address public accountability concerns.
- **Carolynne Muncer** takes us on her journey to the land of formative assessment at the high school level. Questioning the tradition that has blocked her and others to respond to the learning needs of students, she takes the assessment road less traveled. After serious reflection, Carolynne talks with her students about ways to improve her assessment practices. Her story is educative and encouraging.
- **Judy Arter** shares her reflections about the purpose and effectiveness of rubrics. She penetrates the obvious and untangles the inherent contradictions in the use of rubrics. Judy's value-added analysis suggests that quality

rubrics must communicate what counts, guide student learning towards important learning targets, and inform instruction.

- **Lorrie Shepherd** positions assessment within the larger social context of quality learning. Her insights are about making meaningful connections and using concepts such as scaffolding and zones of proximal development to support learning and to guide children's learning. Lorrie discusses four practical and specific strategies for forming and informing learning—assessing prior knowledge, feedback, teaching for transfer, and self assessment. Her article demonstrates the value of creating an authentic learning culture that embodies the best of all assessment practices.

—Clay Lafleur



We are also concerned about the direction that this form of assessment and accountability could lead to, as evidenced by watching other education systems, particularly in the U.S. Will this lead to teachers wanting to transfer to schools in wealthy neighbourhoods where students do better on the tests? Some U.S. schools have funding tied to student test scores, with schools that do better on the tests getting more funding. While such plans for using the EQAO tests have not been discussed widely and publicly in Ontario, we remain cautious. The education policies imposed under the previous Conservative government are still very fresh in teachers' minds. While some of the policies did not need to have a negative impact, the manner in which they were imposed turned the education system on its head. Leithwood (2006) notes that working conditions and the motives attributed to them make a difference.

Overemphasis on EQAO Assessments

The testing and accountability policies that have been widely implemented in many education systems over the past 10 years have changed the culture of some schools, increased competition among teachers in other schools, and in some jurisdictions—such as Ontario—have seriously undermined the status of the profession in the public's mind.

Leithwood also makes it clear that teachers' working

conditions as well as how teachers' feel about the working conditions makes a difference to student achievement. ETFO members are concerned with students and the education system.

One challenge is that an attempt is being made to accomplish too many goals with one set of tests, goals that the tests were not designed to meet. This was identified in the external review of EQAO conducted by researchers at OISE/UT (Wolfe, Childs, & Elgie, 2004). The EQAO assessments are used to evaluate students, schools, district school boards, and the whole provincial education system. Using one assessment for all purposes is not appropriate.

Many teachers and students are concerned that the emphasis on testing takes away time and energy from other areas of the curriculum and results in a narrow approach to the delivery of the curriculum areas that are assessed. Teachers feel pressure to spend increasing amounts of time on the expectations that can be assessed in a pencil and paper assessment and decreasing time on the expectations that require more hands-on learning and therefore are not on the assessment. Some feel pressure to complete the entire curriculum before the administration of the EQAO assessment and therefore have decreased time for student exploration, investigation, and focus on teaching for deeper understanding as they rush to complete all the expectations in language and mathematics before the middle of May. With almost 4,000 expectations in the Ontario elementary curriculum, teachers find that they cannot address all

areas of the curriculum. Science and technology proves to be the most problematic, with almost 80 per cent of public elementary teachers stating they cannot teach all of the expectations (McAdie, 2004). The EQAO tests, with their focus on only a portion of the curriculum, are at least partly to blame for areas of the curriculum being de-emphasized. A letter to the editor of the *London Free Press*, by a student who has been through the tests in Grade 3 and 6 states:

"Our teacher spent the whole year worked up about the tests. Science and history classes were cut short so there would be more time to work on math and English. During the two-week period leading up to the tests, we only had one science class, no history or geography. This meant that the next year we were way behind in these subjects.... What the board needs is a fairer way to evaluate students without putting so much unnecessary stress on both students and teachers." (Wilson, 2005)

Standardized testing contributes to what has been referred to as the de-professionalization of teaching, removing one aspect of professional autonomy from teachers. When the EQAO tests were introduced, the previous Conservative government made it clear that they distrusted teachers—almost demonizing them. The media picked up this lack of respect. Teachers felt demoralized. Teachers felt that their professional judgment was not valued or trusted. Graham, Cameron, and Neu (2004) note that such testing is part of a larger movement to centralize control over education.

Throughout North America, and beyond, school jurisdictions have introduced standardized testing programmes. We see this as part of an increasing movement towards the governance of education from a distance, using funding mechanisms and other administrative techniques. (p. 297)

While the current Ontario Liberal government has hard to support teachers and promote professionalism and the critical role they play in developing knowledgeable citizens, it has also put more weight and importance in EQAO scores than ever.

Is there anything good about the EQAO assessments? One positive aspect of the EQAO tests is that they get more people talking about education, teaching, and learning. However, there are much better ways to stimulate this discussion. For example, the Composite Learning Index, with the first year's results just released by the Canadian Council on Learning, provides a much more comprehensive view of lifelong learning, including a wide variety of measures. The CLI allows communities to look at their total efforts to support lifelong and life-wide learning.

EQAO has started to provide teachers with an item analysis of the tests, showing teachers where they might be able to improve their instruction to ensure more students are learning the needed aspects of at least numeracy and literacy. While this type of assessment for learning is valuable to teachers, we do not need large-scale assessments for teachers as well as students to be able to learn from assessments. Teachers can use data they collect every day with their students to enable this type of analysis and discussion. What teachers really need is time to

Surveys on Teachers' Concerns

The Ontario College of Teachers:

Teachers have been surveyed for the last three years, and each years' results show a strong opposition to the EQAO tests.

2005:

- 71% indicated that standardized testing was the least effective reform for improving student learning (COMPAS, 2005).

2004:

- 8% agreed that the EQAO tests should be used to evaluate performance of individual schools (COMPAS, 2004).

2003:

- 95% stated that the tests were not a good way to track teacher performance,
- 92% stated they are not a good measure of school success,
- 90% stated they do not improve student learning,
- 88% stated they are not a good way to track student success, and
- 85% stated the tests demoralize and de-motivate students (COMPAS, 2003).

Elementary Teachers' Federation of Ontario:

Members were surveyed as part of Teaching for Deep Understanding, a joint project with OISE/UT.

2004:

- 6% agreed that the EQAO tests are valuable for assessing students' progress,
- 12% agree that the tests have a positive impact on teaching, and
- 4% agreed that they improve communication with parents (McAdie, 2004).

Earlier survey work with ETFO members also shows concern with the EQAO tests.

2000:

- 56% indicated that the tests have had a negative impact on the performance of their role as teachers. Of those members, 65% (primary grade teachers) and 60% (junior grade teachers) were more likely to express this concern over province-wide testing (ETFO 2000).

work with colleagues to review student work, discuss misconceptions, and explore instructional strategies.

Lorna Earl (2004) talks of assessment of learning, for learning, and as learning. EQAO falls primarily into the assessment of learning category; it is purely summative. It does not provide teachers or students with useable information since the test happens near the end of the year and the results are not provided until into the next year. In Earl's lexicon, then, it is not assessment for learning. It cannot shape the teachers' instruction until the following year.

The main purpose of the EQAO assessments, then, is as a measure of accountability. While we must not shy away from accountability, this is not the way to do it.

Where Does That Leave Student Assessment?

Teachers use a variety of methods to assess students and to communicate that assessment data to both students and parents—direct observation, interviews, conferences, short written and oral tests, performance tasks of varying complexity, portfolios and collections of student work, and journals and learning logs. Teachers also involve students in the assessment process. They encourage students to become "active, engaged, and critical assessors" (Earl, 2003, p. 25) as they monitor their own learning. With smaller classes, better resources, time to collaborate with peers, more professional development, and even more preparation time, the majority of teachers are working to perfect their teaching and assessment skills.

Teachers are being encouraged to work in professional communities to further their own learning. A great number of teachers take courses in the summer to further their understanding of all aspects of teaching, including classroom assessment. All of these activities demonstrate in concrete ways that teachers are accountable.

Teachers and teacher unions do not believe that the EQAO assessments are valuable. However, clarifying the goals of the assessments would help to improve the situation. For example, we know that such assessments are not sensitive to many school factors, such as socio-economic status. Therefore, we believe the school results should not be publicly reported.

This government was elected on the promise of ensuring that 75 per cent of students reach level 3 or 4 on the EQAO tests by age 12. Ensuring that more students achieve at a higher level is a laudable goal. But tying the goal to EQAO scores is not. "The drive to 75," as it has been labeled by at least one board, diverts attention from other issues that need to be addressed—what should be taught, what resources need to be provided, and what we need to do to ensure all students are accommodated.

References

- Amrein, A.L. & Berliner, D.C. (2002). *An analysis of some unintended and negative consequences of high-stakes testing*. Education Policy Research Unit, Education Policy Studies Laboratory, Arizona State University. www.asu.edu/educ/eps/EPRU/documents/EPSL-0211-125-EPRU.pdf.
- COMPAS. (2005). *State-of-the-teaching-profession: A COMPAS report to the Ontario College of Teachers*. www.compas.ca.
- COMPAS. (2004). *Annual state of education survey for the Ontario College of Teachers: A COMPAS special report on teacher and public perceptions*. www.compas.ca.
- COMPAS. (2003). *The state of the teaching profession in Ontario. A report to the Ontario College of Teachers*. www.compas.ca.
- Canadian Council on Learning. (2006). The 2006 Composite Learning Index, available online at <http://www.ccl-cca.ca/CCL/Reports/CompositeLearningIndex2/2006Results.htm?Language=EN>
- Earl, L. (2004). Classroom assessment for deep understanding. In K. Leithwood, P. McAdie, N. Bascia, & A. Rodrigue (Eds.), *Teaching for deep understanding. Towards the Ontario curriculum that we need*. Toronto: Elementary Teachers' Federation of Ontario.
- Earl, L. (2003). *Assessment as learning. Using classroom assessment to maximize student learning*. Thousand Oaks, CA: Corwin Press.
- Elementary Teachers' Federation of Ontario. (2000). *Report from the 2000 Survey of ETFO Members*. Toronto: ETFO.
- Froese-Germain, B. (1999). *Standardized testing: Undermining equity in education*. Ottawa: Canadian Teachers' Federation.
- Graham, C., & Neu, D. (2004). Standardized testing and the construction of governable persons. *Journal of Curriculum Studies*, 36(3), 295–319.
- Kohn, A. (2000). *The case against standardized testing. Raising the scores, ruining the schools*. Portsmouth, NH: Heinemann.
- Leithwood, K. (2006). *Teacher working conditions that matter: Evidence for change*. Toronto: Elementary Teachers' Federation of Ontario.
- McAdie, P. (2004). What Teachers Think. In K. Leithwood, P. McAdie, N. Bascia, & A. Rodrigue (Eds.), *Teaching for deep understanding. Towards the Ontario curriculum that we need*, Toronto: Elementary Teachers' Federation of Ontario.
- Wilson, A.T. (2005). Letter to the editor: EQAO tests stress students, teachers. *London Free Press*, November 7.
- Wolfe, R., Childs, R., & Elgie, S. (2004). *Final report of the external evaluation of EQAO's assessment process*. Toronto: OISE/UT.



PAT McADIE is a research officer with the Elementary Teachers' Federation of Ontario. Her work for the past 26 years has covered a diverse number of issues, including pay equity, gender, accountability, collective bargaining, privatization of education, and curriculum reform. She writes and presents workshops on education issues, recently conducting workshops with the South African Democratic Teachers Union on building research capacity in teacher unions.



RUTH DAWSON is on staff at the Elementary Teachers' Federation of Ontario working in professional development. She has also worked as an Education Officer with the Ministry of Education and as Co-ordinator of Elementary Mathematics, Assessment and Evaluation with the Halton District School Board. She has developed and presented workshops on elementary mathematics, assessment, evaluation and reporting, and is currently co-writing a teacher resource on classroom assessment practices.

Reducing Bias in Classroom Assessment and Evaluation

■ Louis Volante

LOUIS VOLANTE uses notions of reliability and validity to create a framework for reducing bias in classroom assessment, and develops a foundation that supports improved instruction and learning.

Fairness in assessment and evaluation is grounded in the belief that all students should be able to demonstrate their learning regardless of their socio-economic status, ethnicity, gender, geographic location, learning style, and/or need for special services. These factors typically receive much attention in relation to large-scale assessment reform, particularly since there is a plethora of research suggesting external testing is often culturally biased (Lee, 2004; Stobart, 2005).

Concern over the technical quality of large-scale assessment has led to increased test surveillance and even spurred the development of complex statistical methods to identify bias such as differential item functioning (DIF), and differential test functioning (DTF). Unfortunately, the healthy preoccupation with the fairness of large-scale assessment measures has not led to corresponding levels of concern about the soundness of the data that matters most of all—teachers' classroom assessments.

To date, very little attention has been directed towards improving the assessment literacy of classroom teachers (Earl, 2003; Stiggins, 2001). This is despite the fact that classroom assessment data forms the basis for pivotal decisions such as grade retention and promotion.

Although there is no such thing as an absolutely fair assessment, teachers have an ethical and professional obligation to minimize sources of bias within their classrooms. Bias in classroom assessment and evaluation may be intrinsic to the assessment measure itself, to the assessment process, to the assessor, to the assessee, or to the interplay among these factors. If we want to be sensitive to the diverse needs and backgrounds of students, teachers need to be wary of these sources of bias and take active steps to reduce their deleterious effects.

One way to approach this lofty goal is to develop assessment measures and procedures that are reliable and valid. Although psychometric considerations are usually the domain of professional test developers, teachers also need to become sufficiently conversant with these issues if they are to promote fair assessment practices within their classrooms.

Reliability and Validity

Reliability refers to the assessment score, not the measure from which the score is generated. In order to reduce bias and the error associated with classroom scores, teachers should use a blind process when grading tests, essays, and other products whenever it is possible. For example, by concealing names and marking separate sections of a test individually, teacher's scores become more consistent and are less likely to be influenced by extraneous factors such as the student's previous performance and/or their cultural background. Blind marking may also reduce the tendency to incorrectly assign higher grades in subjects typically viewed as the relative strength of boys (i.e., math and science), girls (i.e., language arts), or particular cultural groups.

Reliability is also improved when teachers provide students with multiple opportunities to demonstrate what they know. For example, unit-culminating tests may be preceded by informal assessments. Just as a doctor gauges blood pressure by taking repeated readings on different days, so too should a teacher utilize multiple data sets before arriving at a decision about a student's relative strengths and weaknesses. In addition to strengthening the accuracy of teachers' evaluative judgments, this approach also provides valuable information to both students and teachers: students are given feedback so that they might intensify their efforts in particular areas; teachers can use the results to inform their future instruction. Thus, enhanced reliability is facilitated through the use of assessment for learning (formative purpose), as opposed to the traditional assessment of learning (summative purpose).

Simply put, validity refers to the accuracy of the inferences generated from the results of an assessment measure. It is

Unfortunately, the healthy preoccupation with the fairness of large-scale assessment measures has not led to corresponding levels of concern about the soundness of the data that matters most of all—teachers' classroom assessments.

important to note that assessments can be reliable without being valid, but never vice versa. For example, a teacher can develop a culminating test on Canadian First Nations that is scored consistently but fails to incorporate different historical interpretations held by certain cultural groups (e.g., First Nations have contested the Bering Strait theory) or may omit content related to significant indigenous groups such as the Metis, Huron, or Iroquois. Thus, depth and breadth of content sampling are two important considerations for sound assessment. Validity is also improved when teachers utilize different methods of assessment that tap multiple intelligences as well as different levels of cognition and learning styles. The use of a single method of assessment will probably further disadvantage students from particular social classes and ethnic groups (Banks, Cookson, Gay, Hawley, Irvine, Nieto, Schofield, & Stephan, 2001). Clearly, developing reliable, valid, and fair assessments are complementary objectives for teachers.

Specific Strategies for Reducing Bias

In addition to improving the reliability and validity of classroom assessments, teachers should consider the following specific strategies for reducing assessment bias:

- Examine how your classroom complements or conflicts with the school experiences and assessment techniques which are familiar to recent immigrants and indigenous peoples.
- Develop test questions and other assessment items that reflect the multicultural and multiethnic composition of your school, district, region, province, and country.
- Utilize gender-neutral terms within tests, quizzes, and other forms of assessment.
- Provide modifications to tests and other assessment measures for students with exceptionalities and those with limited proficiency in English.
- Take steps to reduce the effects of test-wiseness on academic achievement—design carefully constructed test items, utilize multiple-formats, provide reviews prior to testing.
- Involve students in the development of evaluation criteria—develop rubrics with the assistance of students.
- Adopt a range of formative and summative assessment strategies that encompass different ways of demonstrating task mastery—speaking, writing, and performing.
- Balance the weight given to different types of traditional and authentic performance-based assessment data when arriving at final course grades.
- Most importantly, reflect on pre-conceived notions that may affect the marks/grades you assign to particular groups of students.

Collectively, these strategies can help create a welcoming environment for all students and reduce bias that often provides an unfair advantage to dominant cultural groups.

Strategies for Fair Classroom Assessments

- Blind marking
- Gender-neutral terms
- Cultural representation
- Accommodations for special populations
- Different assessment methods
- Multiple data-sets
- Reflection on pre-conceived notions

Professional Development

Reducing bias in classroom assessment and evaluation naturally begins in faculties of education. Unfortunately, students often report poor preparation in this area (Campbell & Evans, 2000). Indeed, there are teacher education programs that offer their students little more than a brush stroke of information on student assessment and evaluation. In many instances, these programs have assumed that their students will develop this knowledge within their curriculum and instruction courses. Clearly, the importance of adopting fair assessment practices should necessitate more focused attention in all teacher education programs within Canada.

Poor pre-service preparation also suggests that practising teachers require ongoing professional development opportunities to develop assessment expertise. The central importance of assessment literacy has recently motivated governments to provide funding and other resources targeted to support the development of sound classroom assessment practices. This type of large initial investment has resulted in culturally sensitive assessment practices in K–12 schools and underscores the value of providing similar large-scale initiatives within a multicultural country such as Canada (Volante & Melahn, 2005). Nevertheless, less grand schemes can also produce fruitful results for individual teachers and schools. For example, one way for a teacher to enhance their assessment knowledge is to simply read an assessment book written specifically for practicing teachers—particularly those sections that address differentiated instruction and diversified assessment. Another approach is to develop learning teams within schools that regularly share ongoing assessment related successes and challenges (Stiggins, 2002). Administrators can facilitate such teams by providing the space available for teacher conferencing, while districts can provide the release time required. Thus, assessment literacy development is an important starting point for the realization of sound and fair assessment techniques within our schools.

... assessment literacy development is an important starting point for the realization of sound and fair assessment techniques within our schools.

Conclusion

I have offered several strategies that can be used to reduce bias in classroom assessment and evaluation. The realization and widespread adoption of these strategies go hand-in-hand with broader efforts to improve classroom assessment. Ironically, the advent and widespread use of large-scale assessment measures is partly a response to the perception that teachers' classroom assessment data are more subjective and less reliable than standardized measures. Whereas policymakers have come to view large-scale measures as the most objective way to evaluate quality in public education (Burger & Krueger, 2003; Jones, 2004), teachers' classroom assessment is often viewed as technically flawed and characterized as "self-interested home cooking" (Popham, 2004). Clearly, one way to combat this perception is to promote the adoption of classroom assessment and evaluation strategies that are psychometrically sound.

If teachers' classroom assessment data can more thoroughly stand up to public scrutiny, perhaps more focused attention can be devoted to alternative conceptions that recognize the value of assessment for learning and assessment as learning (the assessment is not graded but acts as a learning tool). Collectively, these two assessment purposes provide a foundation for authentic teaching and learning that respects student diversity.

References

- Banks, J.A., Cookson, P., Gay, G., Hawley, W.D., Irvine, J.J., Nieto, S., Schofield, J.W., Stephan, W.G. (2001). Diversity within unity: Essential principles for teaching and learning in a multicultural society. *Phi Delta Kappan*, 83(3), 196-198.

- Burger, J.M., & Krueger, M. (2003). A balanced approach to high-stakes achievement testing: An analysis of the literature with policy implications. *International Electronic Journal for Leadership in Learning*, 7(4). <http://www.ucalgary.ca/~iejll>.
- Campbell, C., & Evans, J.A. (2000). Investigation of preservice teachers' classroom assessment practices during student teaching. *Journal of Educational Research*, 93(6), 350-55.
- Earl, L. (2003). *Assessment as learning*. Thousand Oaks, CA: Corwin.
- Jones, K. (2004). A balanced school accountability model: An alternative to high-stakes testing. *Phi Delta Kappan*, 85(8), 584-590.
- Lee, J. (2004). Multiple facets of inequity in racial and ethnic achievement gaps. *Peabody Journal of Education*, 79(2), 51-73.
- Popham, W.J. (2004). All about accountability/Why assessment illiteracy is professional suicide. *Educational Leadership*, 62(1), 82-83.
- Stiggins, R. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan*, 83(10), 758-65.
- Stiggins, R. (2001). The unfulfilled promise of classroom assessment. *Educational Measurement: Issues and Practice*, 20(3), 5-14.
- Stobart, G. (2005). Fairness in multicultural student assessment systems. *Assessment in Education*, 12(3), 275-287.
- Volante, L., & Melahn, C. (2005). Promoting assessment literacy in teachers: Lessons from the Hawaii School Assessment Liaison Program. *Pacific Educational Research Journal*, 13, 19-34.



LOUIS VOLANTE has held faculty appointments at Concordia University and the University of Hawaii. He has also lectured at OISE/UT, where he obtained his Ph.D. Currently, he is an Assistant Professor in the Faculty of Education, Brock University, teaching assessment and evaluation courses at both the pre-service and graduate level. His research interests include classroom and large-scale assessment, assessment policy, educational program evaluation, teacher education and development, and educational change. He shares with Lorna Earl the 2005 R.W.B. Jackson Award for the most outstanding English language journal article.



With THANKS

HEATHER BERKELEY has captured an audience of educators for more than 23 years through her work with OISE Press, *Orbit*, newsletters, web sites, annual reports, and more. She is influential in surprising and ingenious ways. Ideas are her thing and she generates a cornucopia of new ones by the minute. Yet she remains staunchly committed to the task at hand. She will be missed as our fearless *Orbit* editor!

Heather Berkeley retired from OISE/UT in May 2006, and joined The Literacy and Numeracy Secretariat of the provincial government, undoubtedly stirring the halls with her country tunes and fashion sensibilities.

ALL ABOARD!

The Land of Formative Assessment Awaits

■ Carolynne Muncer

CAROLYNNE MUNCER takes us on her journey to the land of formative assessment at the high school level. Questioning the tradition that has blocked her and others to respond to the learning needs of students, she takes the assessment road less traveled.

How can it be that the more we learn about our practice, the more incompetent we feel? How can it be that the smarter we get, the harder our job becomes? As I dove into the reams of current literature and research on assessment for my Master's program last summer, these thoughts created great angst and revelation. I thought, "I've been a teacher (K-12) for 15 years! How could I have been so out of touch with 'assessment reality'? What have I done to all those kids I've worked with?"

I took stock of my successes and failures as a teacher, and accepted that there were many things I knew I should be doing, should not be doing, and should be trying in my classes. I re-read Lorna Earl's article "Classroom Assessment for Deep Understanding: Shifting from Assessment of Learning to Assessment for Learning and Assessment as Learning" (2003). Earl asks three simple questions: What is our purpose? What are we trying to accomplish? What is assessment for? Earl boldly goes where many educators fear to tread; to the land of uncertainty, discomfort, and promise—the land of formative assessment.

As I embarked upon my graduate studies with Earl, I realized what she was questioning was the same thing Stiggins (2004) and Leahy, Lyon, Thompson, and Wiliam (2005) questioned in their articles on assessment—tradition. Why do we continue to follow assessment practices that no longer meet the needs of learners or teachers? Because it's always been done that way? Or is it because it's easier to follow the rest of the herd over the cliff than to forge new ground? Are we so caught up in "quality control" (grades, percentages, statistical comparisons) that we've lost sight of "quality assurance" (criteria for success, guidance for instruction, self/peer assessment) for our students? (Leahy, Lyon et al., 2005, p. 19).

I had been on this road before. As an elementary teacher, I tried to weave many formative assessment practices into my lessons. It was easier because the K-7 curricula naturally lends itself to these methods of assessment. However, once I made the leap to the secondary level, I found myself lost in a sea of summative assessment. There was so much content to cover and

so little time to do it in. I continued to share criteria for success, provided reams of feedback on how to improve, and tried to provide opportunities for my students to peer and self assess. I fought to keep up with the amount of content that needed to be covered—I taught, I tested, and I moved on. I knew I wasn't meeting the needs of all my learners, but I tried to meet the needs of the school district and the ministry: providing marks, data, and accountability. As such, secondary assessment practices became the crux of my angst.

I was left with burning questions which forced me to analyze my own practice closely. As an elementary-trained teacher, I was taught a bit about formative assessment and how to use it in primary and intermediate grades. I also knew from reading the article "Classroom Assessment Minute by Minute, Day by Day" (Leahy, Lyon et al. 2005) that there were five broad assessment strategies that melded nicely with Earl's work. What I liked about these strategies was that they could be applied to all grades and content areas, *even* at the secondary level:

- ensuring students understand learning intentions/criteria for success;
- creating effective discussions, questions, and assignments;
- providing useful feedback;
- encouraging students to take ownership of their learning; and
- using peer assessment and feedback as a teaching resource.

These assessment gurus made me wonder: Were these strategies still a consistent part of my practice at the secondary level; particularly in my senior English classes? Did I use any of them? If so, did I use them well? Did they improve my students' learning? And, if I was not using all of the strategies, which ones were they, and why not?

To get an honest answer, I went straight to the source—my students. I asked for volunteers to critique a paper I was writing on assessment and for them to assess *my* practice. Opening myself up to the "shoot from the hip" criticism from my toughest critics was scary. I now understood what it felt like to be a student—absolutely vulnerable!

My students' feedback provided answers to my questions and unexpected insights into my instruction. In short—I was using formative assessment more than I thought. Overall, my students felt our discussions and assignments were meaningful and effective; the peer assessments and self evaluations I *forced* them to do, though painful, had paid off and were the most effective in improving their writing. Intentions and criteria for learning were always clear and present in my lessons.

What I wasn't expecting to learn was that my feedback (which took hours) was occasionally useful, but more often than not "useless." Ouch! They said that pointing out what they do well, rather than what they need to work on, was a much more powerful tool to improve their writing. This is sometimes hard to do; especially when students' skills are well below grade level. But it is something I am working hard at to improve. They also indicated how helpful "one-on-one" time with me was towards improving their work. This is something all teachers struggle with; too many students and not enough teacher to go around. In response to this feedback, I now use more peer assessments so that I can work with individuals or small groups on areas of need. I have also set up "experts" in the class for students to go to. As with all writers, my students have areas of weakness and strength so I decided to capitalize on their strengths. Each week, a student is designated to be one of the following experts: the Transition Trickster, the Thesis Master, the Structural Engineer, the Wizard of Rhetoric, and the Quote Integration Specialist. Using peer assessment forces students to take ownership of their learning, and they see me as a coach rather than the "be-all, know-all" of content. Students still prefer my time, but they are more willing to use their peers as resources to improve their writing.

Time is an essential ingredient in the learning process. It is also our enemy in the classroom. It forces us to rush through curriculum, leaving little time to actually teach it. Students said that sometimes my pace was too quick and that I tended to teach to the upper end of the class. This observation stunned me. Maybe this happened because they (the "smarter" kids) are openly engaged in learning and tend to dominate a discussion or lesson. That is not to say I ignore those that are not. I try very hard to draw them in, encourage participation and to create a learning environment where all students feel safe enough to take risks. Unfortunately, engaging at-risk learners is often a challenge. Many of my students miss a great deal of school and, as such, their skills are so weak that it is a great challenge to modify/adapt the curriculum to meet their needs, while still preparing them for the provincial exam.

My responses to this feedback are not excuses; rather they are facts. They are "my facts" to deal with and (thanks to my students) I am much more aware of my tendency to leave some students "behind." I pay closer attention to those "blank looks" which now provide me with ample cues to reassess my instruction. The questions I posed were answered: Did I use formative assessment in my classes? Yes, but not enough. Am I alone on this journey? Probably not! Finally the revelation.

Overall, my students' feedback was great professional development and a very humbling experience. I learned that even though my intentions were good they often hindered rather than helped my students achieve their full potential. I became conscious of two important things. First, as teachers, we must be willing to take ownership of our failures and weaknesses, recognize our strengths and build on them, ask for support when needed, and share our expertise with colleagues in order to improve the way we teach and assess students. It may be more

work, but it's worth it. We ask students to do these things every day in our class, so why should we be any different?

Second, I realized exactly what Earl is driving at. Assessment of student learning and our own teaching practice is a balancing act. Certainly there is a clear purpose for summative assessment. It sets standards and establishes baselines for accountability and achievement. It also raises questions about equity in our system. However, without assessment for learning, our instruction does not address what students know, how they use what they know, and how we can highlight their strengths and weaknesses. Moreover, without assessment as learning (the most uncomfortable leg of this journey), our students will never be fully aware of their own unique learning styles or their ability to be a "critical connector" between assessment and deep learning (Earl, 2003, p. 4).

I think that most teachers and schools are all over the "assessment map." Those who want to know more and do better find their way by hook or by crook. There are those old dogs who, even if given the bone (time, resources, and support), will refuse to learn these formative assessment "tricks." Finally and unfortunately, there are those who profess to "know it all already" and do not recognize their own incompetence.

My trip with Earl and her fellow assessment gurus has reacquainted me with assessment knowledge I forgot I possessed, and provided me with much more to work with. I no longer feel incompetent. However, I am still overwhelmed by the amount of work secondary schools have to do to ensure our instruction and forms of assessment meet our students' needs and move them forward. If formative assessment is going to be a seamless part of the learning process in every classroom and school, we must be willing to embark on this challenging journey and boldly bring our students to this land of promise (Earl, p.7).

References

- Leahy, S., Lyon, C., Thompson, M., & Wiliam, D. (2005). Classroom assessment minute by minute, day by day. *Educational Leadership*, 63(3), 19-24.
- Earl, L. (2003). *Assessment as learning: Using classroom assessment to maximize student learning*. Thousand Oaks, CA: Corwin Press.
- Stiggins, R. (2004). New assessment beliefs for a new school mission. *Phi Delta Kappan*, 86 (1), 22-26.



Carolynne Muncer has 15 years of teaching experience at both the elementary and secondary level. She has taught in a one-room school and in remote First Nations communities. For the last five years, Carolynne has been a career counselor and a senior English teacher at a small high school in British Columbia. This school year, she began her formal administrative journey as an elementary principal.

What Do We Want Rubrics to Do For Us and Our Students?

■ Judy Arter

JUDY ARTER shares her reflections about the purpose and effectiveness of rubrics. She penetrates the obvious and untangles the inherent contradictions in their use.

Lately I've been considering the question of what we want rubrics (those written-down criteria for judging the quality of student work or performance) to do for us and our students. Are they simply a mechanism for giving a mark or grade (and letting students know what is required to get a good mark or grade), or can they be more?

I began thinking about this several years ago when I read about a third grade teacher who encouraged her students to develop the scoring guides for some of the work they produced. In one instance, the teacher asked her students to make a poster to demonstrate their understanding of penguins. The students developed the following list as the criteria for judging the quality of their posters:

- The poster is neat; we can read it.
- The poster has at least three colours.
- The poster has at least five facts about penguins.
- Facts about penguins are correct.
- The poster has both pictures and words.
- The poster is done on time.

Students used these criteria to help each other make their posters better. Finally, the teacher assigned a mark for the poster based on the criteria and put the mark in her record book to count towards a final grade in science.

On the surface it's hard to argue with this procedure. The teacher involved students at all steps of the assessment process so that they understood exactly what it took to get a good mark. The students self and peer assessed and were encouraged to revise their work before the final mark. The teacher gave students control over the circumstances of their own success to improve student motivation and ability to do well. So, where's the hitch?

What is the Mark Evidence of?

There are at least three different student learning outcomes mixed up in the mark:

- work habits (getting the poster done on time);
- art (poster neatness and colourfulness); and
- science (understanding penguins).

This confounding is a problem. In a standards-based system, how would this single mark get transformed into a judgment of mastery on standards? Mastery of which standards?

At least, to be meaningful and helpful, there should be separate marks for content understanding, work habits, and art. These could then be placed in their respective portions of the grade book to count towards a final judgment of mastery in those respective areas.

Are the Stated Criteria Really Important?

For example, why *three colours*? Under the right circumstances, couldn't an effective poster be black and white, red and green, or blue and yellow? What message does *three colours* send to students about the nature of a quality poster? "Oh, I get it, I always need to use three colours when I make a poster." What would the teacher re-teach if the students failed to use three colours? "Next time have three colours." Where in our content standards does it say that students need to learn that a good poster has three colours?

There's a similar argument to be made for *five facts*. Wouldn't three really interesting and salient facts be better than five peripheral facts? Most times quantity is not the same as quality. What message does *five facts* send to students about what they were supposed to learn from the project? What would the teacher re-teach if the students failed to have five facts? What would teachers bring forward to the next project to extend learning over time?

What is the Important Learning Intended?

This question goes even deeper: what is the important learning to be achieved here? What is *three colours* an indicator of? Might the real criteria for designing an effective poster be that the poster draws the viewer in; makes the viewer want to look at what we have to say; and having hooked the viewer, the design enhances understanding?

What we want learners to realize is that sometimes these criteria can be accomplished by the use of colour; other times they can be produced by size, layout, or other things.

Instead of a poster *needs to have three colours*, a better message about quality would be sent to students if the criterion were your poster *should be designed to draw in the viewer so that she wants to look at your work and your poster should be designed to help the viewer understand what you want to say*. (Would this also give you better information for tracking student

Do we want rubrics to simply provide a mechanism for allowing students to get a good mark or grade, or do we want rubrics to define and communicate the nature of quality as it relates to important and well-defined learning targets?

performance and planning the next steps in instruction?) Or better yet, instead of narrowing it down to a poster (after all, where in our content standards does it say that students are to learn how to make a good poster?) we broaden the concept to include all products: *Anything you create should be designed to (a) make the viewer want to consider your work and (b) enhance viewer understanding.* Think of the added scope and flexibility this would give the teacher and students to develop, over time and across projects, student understanding of the concept of presentation?

Thus the poster would become one of several contexts in which to focus student attention on keys to effective presentation. Dioramas, visual aids in an oral presentation, and other projects could also be used to develop that same concept over time. Then, if the teacher kept her grade book by standard, she could keep track of student development on mastery of the concept of presentation. (I am assuming here that effective presentation is represented in content standards.)

Similarly, what is *five facts about penguins* an indicator of? To which standard does this criterion contribute? The teacher references it to science, but which science standard? It seems that learning about penguins per se isn't even the goal. The real goal might be something like *students can choose the most important or most interesting facts for a given audience—or—students will understand how habitats affect animal appearance and behaviour.* The former might be a language arts goal and the latter would be a science goal.

If the former goal is the intended learning, a better message about quality would be sent to students if the criterion, instead of *five facts*, was *the facts you choose should be those you think would be of most interest to other students in the class and choose details that make other students interested in what you have to say.* Changing to this criterion would tie the assessment directly to an important learning outcome, adding the scope needed for the teacher to build, over time, student mastery of the ability to write for different audiences and purposes.

What Do We Want Rubrics to Do for Us?

I think all these considerations relate back to the question of why we want to use rubrics in the first place. Do we want rubrics to simply provide a mechanism for allowing students to get a good mark or grade, or do we want rubrics to define and communicate the nature of quality as it relates to important and well-defined learning targets?

If the goal is simply to communicate to students what they need to do to get a good mark, anything can go in the list of criteria: length, neatness, use of vocabulary, use of a five-paragraph theme format, having a paragraph for each main point.... Students can hit any target we clearly set up for them. But where does it take them over time? If the goal is to use the rubric to help students improve their performance over time on important learning targets, then we need to decide:

- What the real learning is—which important learning targets are to be developed through each project or task; the task needs to align with important learning targets
- What the true indicators are of a quality performance on the part of students; if we involve students in developing a rubric for judging quality, we need to be prepared to lead them, through the use of examples and discussion, to the *real* criteria for success; the rubric needs to align with important learning targets

Rubrics can be powerful tools to define the essence of quality so that student involvement in using the rubric to self and peer assess and revise work leads to real learning on important targets—if we keep two questions in mind: What message is this rubric sending to students about what counts? and What information will this rubric give me and my students to help us track student learning towards important learning targets and plan relevant next steps in instruction?



Judy Arter is an Education Solutions Consultant with ETS Assessment Training Institute in Portland, Oregon. She is a co-author of *Scoring Rubrics in the Classroom: Using Performance Criteria for Assessing and Improving Student Performance* (2001), *Assessment FOR Learning: An Action Guide for School Leaders* (2005), *Classroom Assessment for Student Learning: Doing It Right—Using It Well* (2004), and *Creating and Recognizing Quality Rubrics* (2006).

Creating Coherent Formative and Summative Assessment Practices

■ Lorrie A. Shepard

LORRIE SHEPARD demonstrates the value of creating an authentic learning culture that embodies the best of all assessment practices, and discusses four practical and specific strategies for forming and informing learning.

New ideas about how to use formative assessment to enhance student learning must confront the ominous presence of summative assessment and grading. Because of the paramount importance of grades, students worry about what will be on the test rather than thinking about learning. They finish assignments to “be done” and don’t expect to use knowledge again. They’ve learned to pretend competence and hide their confusion from the teacher. The older the students, the more ingrained the negative presumption is that assessment is all about grading. Teachers likewise worry about having a point system that can be defended to students and parents, rather than seeing assessment as a means to illuminate learning goals and further deep understanding.

Recent research has shown that formative assessment can be a powerful tool for moving learning forward (Black & Wiliam, 1998). But to realize the full potential of formative assessment requires a fundamental shift in how learning goals are represented and how assessment insights are used. To create a learning culture (Shepard, 2000) instead of a grading culture, students have to know why what they are studying is important and how it connects to things that they’ve studied before. They need to gain experience talking about their thinking and trust that their teacher’s questions are intended to help them learn.

1 Using Sociocultural Theory in Classrooms

According to sociocultural theory, children develop cognitive abilities through social interactions that let them try out language and practice their reasoning. Instead of being born with an IQ of a certain size, children become “smart” through what Barbara Rogoff (1990) calls an “apprenticeship in thinking.” In various learning contexts—talking at the dinner table, helping in the kitchen, doing math in classrooms—learners have both expert models and supports from adults or peers to enable them to participate in that activity. This process of providing support to help the learner attempt and then master increasingly complex tasks is called *scaffolding*.

Scaffolding, communities of practice, and classroom discourse reforms are all based on sociocultural theory and Vygotsky’s (1978) idea that learners take into their minds ways of thinking that are first practised in social interaction. As Ann Brown (1994) explained, higher cognitive processes are in part “internalized dialogue.” Mathematics and science reforms have students explaining their reasoning to support deeper conceptual understanding but also to help them learn how to reason from evidence. In literacy instruction, classroom routines like “author’s chair” help students think about ways to improve their story or essay, but also help them try out the roles of author and critic. An important aspect of sociocultural theory is that it brings together motivation and cognitive development. At the same time they are developing specific skills, children are also taking on the identity of an author, a scientist, and a capable learner.

The steps in formative assessment are closely tied to the idea of scaffolding and Vygotsky’s zone of proximal development (ZPD). The ZPD is the space between the student’s actual level of mastery, where he can perform independently, and the level of potential development where he can perform with help from adults or in collaboration with peers. Similarly, Sadler (1989) offered a model of formative assessment where the learner has an understanding of the goal, is aware of his or her present level of performance, and then takes action to close the gap.

2 Specific Formative Assessment Strategies

In addition to the general notion that formative assessment is used to “form” and “inform” learning, there are specific formative strategies that have a solid foundation in learning theory and cognitive research.

Prior knowledge assessment—Learning occurs when students make sense of new experiences and ideas and connect them to what they already know. Effective instructional strategies build on what students already know. However, this means knowing more than students’ scores on a pretest. Engaging with students’ relevant knowledge requires a *qualitative* understanding of students’ existing conceptions. What words do they use to describe a phenomenon? How is “coherence” in writing like “cohesion” in chemistry? What have they learned about this topic in school? What do they know about it in real life?

Prior knowledge assessments could include a formal pretest or quiz, especially for older students, but more often students can

Ways Informative Assessments Inform Learning

• Prior Knowledge Assessment

When students share what they know about a subject when starting a study unit, they learn that it helps to think about what they already know when trying to solve a new problem.

• Feedback

Focusing on features of the task, such as how to improve in relation to the standards and emphasis on learning goals, results in positive learning outcomes. Lavishing non-specific praise or making normative comparisons undermines learning.

• Teaching for Transfer

Talking about how and whether students understand the principle can let them generalize from one application to the next. A wide range of questioning in the classroom helps make students more comfortable in the face of extensions, applications, reformulations, and connections.

• Self assessment

Students must think about their work, apply criteria in the context of their own work, and monitor their own learning—increasing responsibility for their own learning.

share what they know at the start of a lesson or unit through instructional conversations. A well-known example is Olge's (1986) K-W-L technique, where teachers ask students to post on a chart what they already *know* (K) and, through discussion, establish what they *want* (W) to learn. At the end of the activity, students discuss and summarize what they have *learned* (L). By participating in this "custom," students become aware that it helps to think about "what I already know when I'm trying to solve a new problem."

Feedback—We know that feedback is essential for learning, but surprisingly, in a comprehensive synthesis of 131 studies, Kluger and DeNisi (1996) found that one-third of studies showed negative effects. In some circumstances, the wrong kind of feedback actually undermines learning. According to Kluger and DeNisi, positive learning outcomes are more likely when feedback focuses on features of the task—such as how the student can improve in relation to the standards—and emphasizes learning goals instead of lavishing non-specific praise or making normative comparisons.

Research on effective feedback brings together findings from both the cognitive and motivational literatures. Happily, they tell

a consistent story. Students aren't born with a certain amount of intrinsic motivation. It develops through social interaction. Students can learn to be more or less intrinsically motivated or to be internally motivated in some settings (often outside of school) and not in others. The same kinds of feedback that support cognitive gains, i.e., focusing on the task and how to improve, also support students' developing sense of mastery and motivation to persist. In contrast, extrinsic rewards, like grades, can reduce intrinsic motivation. In some studies, for example, rewards increase the reinforced behaviour during the study, but when the study is over, students actually do less of the desired behaviour than before the study began.

Teaching for transfer—A goal of learning is for students to be able to apply their knowledge in new contexts. Yet, this is often difficult for novices because their knowledge is fragmented and tends to focus on the surface features of problems. By contrast, experts have deeper and more principled understanding. Just as prior knowledge techniques are intended to help students "be more metacognitive" about their use of knowledge resources, teaching for transfer also makes metacognitive goals explicit by asking students to talk about how and whether they understand the principle that can let them generalize from one application to the next. "What do we already know about fractions that can help us understand decimals?" or "How is learning about ratios and proportions the same as—and different from—learning about fractions?"

Relevant to the idea of transfer is students' often narrow conception of what makes a test *fair*. McClymer and Knoles (1992) tell the true story of a college professor who taught students problems like this:

A boy drops a water balloon from a window. If it takes 0.8 seconds to strike his erstwhile friend, who is 5 feet tall, how high is the window?

On the exam, the professor asked this problem:

Someone walking along the edge of a pit accidentally kicks in a small stone, which falls to the bottom in 2.3 seconds. How deep is the pit?

One student thought this was terribly unfair because they had never done any "pit problems" in class. While it would be unfair to surprise students on a test with the first extensions and applications of what they had learned before, it is not unfair to ask for applications that are within the range of transfer tasks that students have already practiced in class. As soon as students show me they've "got it," a new question always follows my congratulatory smile: "That's great, now let me ask you about that in a different way." When this wider range of questioning is the norm in classrooms, it becomes appropriate to expect extensions, applications, reformulations, and connections on summative examinations.

Self assessment—Student self assessment is not about saving teachers from the work of grading papers. In addition to acquiring specific knowledge and skills, becoming competent in a field of study means learning and internalizing the standards by which others in the discipline judge performance. The process of self assessment requires students to think about and apply criteria in the context of their own work. In so doing, they make sense of and come to understand what the criteria mean in a deeper way than if they merely read posted grading criteria. This kind of supported practice, where students learn strategies to monitor their own learning, helps develop students' metacognitive abilities. At the same time, self critique can increase students' responsibility for their own learning and make the relationship between teacher and student more collaborative.

Ultimately the habit of self assessment leads to the self-monitoring of performance, which is the goal of instructional scaffolding as well as being the goal of Sadler's (1989) formative assessment model. In Klenowski's (1995) study in England and Australia, students who participated in self evaluation became more interested in the criteria and substantive feedback than in the grade they received. Students also reported that they had to be more honest about their work because they had to be prepared to defend their opinions in terms of the evidence.

3 Rich Embodiment of Learning Goals

Assessment cannot promote learning if it is based on tasks or questions that divert attention from the real goals of instruction. Historically, traditional tests have often misdirected instruction, if they focused on what was easiest to measure instead of what was important to learn. Similarly, all of the nurturing and insightful formative assessment processes thus far envisioned will be undone if learning activities encourage rote learning (perhaps like the "pit problem" above) or if grading practices misdirect effort. To be mutually supportive, formative and summative assessments must be conceptually aligned. They must do a good job of representing important learning goals and should use the same range of tasks and problem types to tap students' understandings. Summative assessments can be thought of as important milestones on the same learning trajectories that undergird formative assessment.

As illustrated by Wiggins and McTighe (1998), devising assessments that embody the standards and goals of instruction is central to good teaching, not just a matter of measuring outcomes. Wiggins and McTighe use a process of backward design that begins with the goals, then asks what would be compelling evidence or demonstrations of learning, and last plans activities that would enable students to perform. With understanding as the goal of instruction, Wiggins and McTighe emphasize the need to spell out what evidence of understanding

Grading that Supports Learning

- Make sure that grades are faithful representations of achievement.
- Use the process of "backward design": begin with the goals, ask what would be compelling evidence or demonstrations of learning, and plan activities that would enable students to perform.
- Spell out what evidence of understanding looks like.
- Focus instructional activities and assessments on rich and authentic tasks.
- Emphasize the value of increasing one's skill and participation in an endeavour.
- Focus on achieved competence: create a learning environment that is intrinsically motivating and engaging for students.
- Allow students to apply what they learned from feedback: review how earlier assignments are relevant; use prior knowledge conversations.
- Grade in terms of mastery standards: use other support systems for students of different abilities; differential pacing for learning and timing of benchmark assessments, identification and assessment of intermediate, attainable goals; and differential scaffolding.
- Soften student worries about grades: give replacement assignments and tests; throw out test scores when learning is verified by later assessments; postpone giving grades; use student self and peer assessment; provide "as if" grades.

would look like. These descriptions of performance in turn propel them to provide students with opportunities to develop and practice these skills.

4 Summative Assessment and Grading

Summative assessment and grading pose a serious threat to the learning purposes avowed for formative assessment. If tests or graded assignments diverge from valued learning goals, students focus only on the graded portion of the curriculum. Therefore keeping instructional activities and assessments focused on rich and authentic tasks and making sure that grades are faithful representations of achievement are the first requirements for a grading system that supports learning.

Just as important as content, however, is the effect of grading practices on motivation and interest. The use of grades as rewards contributes to what Lave and Wenger (1991) termed the "commodification of learning" (p. 112). When there is no "cultural value" for increasing one's skill and participation in an endeavour, the only reason to participate is to obtain surface knowledge that can be displayed for evaluation. As explained by Stipek (1996), rewards work to decrease intrinsic motivation when they are perceived as controlling and are not directly related to successful performance. Consistent with the positive findings from research on feedback, rewards or praise that convey *positive information about competence* are more likely to increase intrinsic motivation.

Thus, both the cognitive and motivation literatures argue for a standards-based grading system that focuses on achieved competence. Yet, this recommendation is inconsistent with data showing that most teachers, at least in the U.S., tend to also consider effort, improvement, ability, and behaviour when determining grades (Brookhart, 1994). Often grading practices reflect what I call "compliance grading" such as points given for turning in homework (that is never assessed) or even giving points for bringing a pencil to class. Teachers also tend to have safety net or peacekeeping mechanisms such as extra-credit points that typically do not connect back to the original learning goals. Note that using effort grades to manage student behaviour is not the same thing as creating a learning environment that is intrinsically motivating and engaging for students.

To create a grading system focused on achievement and at the same time allow for the learning potential of formative assessment, teachers may literally have to abandon micro-managing point systems. To help students stay focused on substantive feedback, teachers may need to postpone giving grades or use only student self and peer assessment and, when needed, provide "as if" grades if students are worried about how they will eventually measure up to standards. Importantly, for formative assessment to contribute to learning it must occur in the midst of learning. That means on formally evaluated pieces that there must be subsequent occasions for students to apply what they learned from feedback. This could be accomplished by either allowing time for revision or making transparent in subsequent units how earlier assignments are relevant (in prior knowledge conversations, for example).

In heterogeneous classrooms, grading in terms of mastery standards should include strategies such as differential pacing for learning and timing of benchmark assessments; identification and assessment of intermediate, attainable goals; and differential scaffolding. If taken seriously, the commitment that grades represent achievement would mean doing away with compliance grading, like giving points for turning in note cards; but as part of the cultural shift that scaffolds the development of students' metacognitive abilities, the benefits of assignments that enable performance on culminating assignments should be made transparent for students. Other ways to soften student worries about grades may provide opportunities for students to demonstrate mastery—e.g., replacement assignments and

replacement tests or throwing out test scores when learning is verified by later assessments.

Ultimately our goal is to develop a learning culture where the normative practices in the classroom make it normal and customary to learn from mistakes, to critique one another so that our reasoning can improve, and to develop in students a clear, internal sense of pride and accomplishment on the occasions when summative demonstrations of achievement are needed.

Note

Portions of this article are adapted from Shepard, L.A. (2006). Classroom assessment. In R.L. Brennan (Ed.), *Educational measurement* (4th ed.). Westport, CT: Greenwood Publishing Group.

References

- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment and Education: Principles, policy and practice*, 5(1), 7–75.
- Brookhart, S.M. (1994). Teachers' grading: Practice and theory. *Applied Measurement in Education*, 7(4), 279–301.
- Brown, A.L. (1994). The advancement of learning. *Educational Researcher*, 23, 4–12.
- Klenowski, V. (1995). Student self-evaluation process in student-centered teaching and learning contexts of Australia and England. *Assessment in Education*, 2, 145–163.
- Kluger, A.N., & DeNisi, A. (1996). The effect of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119, 254–284.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.
- McClymer, J.F., & Knoles, L.Z. (1992). Ersatz learning, inauthentic testing. *Journal on Excellence in College Teaching*, 3, 33–50.
- Ogle, D.M. (1986). K–W–L: A teaching model that develops active reading of expository text. *The Reading Teacher*, 39(6), 564–570.
- Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. New York: Oxford University Press.
- Sadler, R. (1989). Formative assessment and the design of instructional assessments. *Instructional Science*, 18, 119–144.
- Shepard, L.A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4–14.
- Stipek, D.J. (1996). Motivation and instruction. In D.C. Berliner & R.C. Calfee (Eds.), *Handbook of Educational Psychology* (pp. 85–113). New York: Simon & Schuster Macmillan.
- Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Wiggins, G., & McTighe, J. (1998). *Understanding by design*. Alexandria, VA: Association for Supervision and Curriculum Development.



LORRIE SHEPARD is professor of Research and Evaluation Methodology and Dean of the School of Education at the University of Colorado at Boulder. Her research focuses on psychometrics and the use and misuse of tests in education settings. Her current work addresses the use of classroom assessment to support teaching and learning.



“S” is for Sam

And most of Sam's favourite things. Sticks the squirrel, story, shell, sun, soccer ball... preschoolers like Sam are discovering the world one word at a time. TVOKids creates lots of exciting ways to introduce young children to letters, phonics, and words so that they can enter school with a positive attitude. Whether it's joining our preschool book club, watching a commercial-free curriculum-based show, or playing any of our games at tvokids.com, learning has never been so much fun!

tvokids
All day, every day.

Gisèle's Big Backyard

Weekdays from 6 am to 3:30 pm

The Space

Weekdays from 3:30 pm to 7 pm

Weekends from 6 am to 4 pm



Now available on your mobile phone. Find us online at www.tvokids.com and on Channel 2 (cable channel may vary in some areas), Bell ExpressVu: Channel 265, Star Choice: Channel 353

Celebrating 100 years of studies in education at the University of Toronto 1907-2007

The 100th Anniversary

Beginning in 2006 the Ontario Institute for Studies in Education of the University of Toronto will host the 100th anniversary of studies in education at the University. The OISE community, which includes faculty, staff, students, alumni and friends, as well as members of the University of Toronto and broader communities will gather together to participate.

The Centennial Book

The Centennial Book, titled *Inspiring Education: Celebrating 100 Years of Studies in Education at the University of Toronto*, retraces the origins of OISE and its antecedents through pictures and stories. Priced at \$24, the book will be available to order online (www.100years.oise.utoronto.ca), in the Education Commons (main floor library) at OISE, 252 Bloor St. W. and at the University of Toronto Bookstore (www.uoftbookstore.com). 10 per cent of book sales will go towards The Centennial Legacy Scholarships. Please join us for the book launch on January 17, 2007.

The Centennial Lecture Series

- February 7, 2007; Isabel Bader Theatre
Teaching Excellence panel discussion
Featuring the five inaugural U of T President's Teaching Award recipients
- March 23, 2007; Education Commons, OISE
Dean's Graduate Student Research Conference
Speaker, Professor Jonathan Jansen
Faculty of Education, University of Pretoria, South Africa
- November 2007; location to be confirmed
2007 Jackson Lecture
Speaker, Professor Rosemary Tannock
Canada Research Chair, Special Education and Adaptive Technology
OISE and Hospital for Sick Children

The Centennial Archival Collection

OISE will showcase its remarkable heritage in a series of archival displays including photographs and artifacts from the past. These will be located in the Education Commons (main floor library) of OISE.

January to February 2007: *Students; Grads*
March to April 2007: *Our Research; Pioneering Women*
May to August 2007: *Curriculum; Buildings*
September to November 2007: *OISE around the World*

The Centennial Open House and Reunion Saturday, May 5, 2007

OISE will open its doors to alumni, retired faculty and staff, the University and our community neighbours to showcase research initiatives in each department and to reconnect alumni with former professors and classmates.

Convocation June 2007 Teacher Recognition Program

OISE will celebrate elementary and secondary school teachers in conjunction with the June 2007 Convocation. Designed to recognize teachers who have had a profound impact on students, the program builds on the idea that educators "stand on the shoulders of great teachers who came before" them.

The Centennial Celebration Party

On Saturday, November 17, 2007 at the Royal Ontario Museum, OISE will host an anniversary party as a wrap-up to the full year of centennial celebrations.

Get involved. We'd like you to be a part of the 100th Anniversary celebration.

www.100years.oise.utoronto.ca

OISE

ONTARIO INSTITUTE FOR STUDIES IN EDUCATION
UNIVERSITY OF TORONTO